Bridging the Evidence Gap in Obesity Prevention: A Framework to Inform Decision Making

ISBN
978-0-309-14989-1

336 pages
8 1/2 x 11
PAPERBACK (2010)

Shiriki K. Kumanyika, Lynn Parker, and Leslie J. Sim, Editors; Committee on an Evidence Framework for Obesity Prevention Decision Making; Institute of Medicine

BRIDGING THE EVIDENCE GAP IN OBESITY PREVENTION

A FRAMEWORK TO INFORM DECISION MAKING

Committee on an Evidence Framework for Obesity Prevention Decision Making
Food and Nutrition Board

Shiriki K. Kumanyika, Lynn Parker, and Leslie J. Sim, *Editors*

INSTITUTE OF MEDICINE
OF THE NATIONAL ACADEMIES

THE NATIONAL ACADEMIES PRESS
Washington, D.C.
www.nap.edu

THE NATIONAL ACADEMIES PRESS 500 Fifth Street, N.W. Washington, DC 20001

NOTICE: The project that is the subject of this report was approved by the Governing Board of the National Research Council, whose members are drawn from the councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. The members of the committee responsible for the report were chosen for their special competences and with regard for appropriate balance.

This study was supported by Tracking No. 20604633 between the National Academy of Sciences and Kaiser Permanente; by Grant No. 61747 with the Robert Wood Johnson Foundation; and by Contract No. 200-2005-13434, Task Order 13 with the Centers for Disease Control and Prevention. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the view of the organizations or agencies that provided support for this project.

Library of Congress Cataloging-in-Publication Data

Bridging the evidence gap in obesity prevention : a framework to inform decision making / Committee on an Evidence Framework for Obesity Prevention Decision Making, Food and Nutrition Board ; Shiriki K. Kumanyika, Lynn Parker, and Leslie J. Sim, editors.
p. ; cm.
ISBN 978-0-309-14989-1 (pbk.)—ISBN 978-0-309-14990-7 (pdf) 1. Obesity—Prevention. 2. Evidence-based medicine. 3. Outcome assessment (Medical care) I. Kumanyika, Shiriki Kinika, 1945- II. Parker, Lynn. III. Sim, Leslie J. IV. Institute of Medicine (U.S.). Committee on an Evidence Framework for Obesity Prevention Decision Making. V. Institute of Medicine (U.S.). Food and Nutrition Board.
[DNLM: 1. Obesity—prevention & control. 2. Decision Making, Organizational. 3. Outcome and Process Assessment (Health Care) 4. Program Evaluation. WD 210 B851 2010]
RA645.O23B75 2010
362.196′398—dc22
2010023392

Additional copies of this report are available from the National Academies Press, 500 Fifth Street, N.W., Lockbox 285, Washington, DC 20055; (800) 624-6242 or (202) 334-3313 (in the Washington metropolitan area); Internet, http://www.nap.edu.

For more information about the Institute of Medicine, visit the IOM home page at: **www.iom.edu.**

Printed in the United States of America

The serpent has been a symbol of long life, healing, and knowledge among almost all cultures and religions since the beginning of recorded history. The serpent adopted as a logotype by the Institute of Medicine is a relief carving from ancient Greece, now held by the Staatliche Museen in Berlin.

Suggested citation: IOM (Institute of Medicine). 2010. *Bridging the Evidence Gap in Obesity Prevention: A Framework to Inform Decision Making*. Washington, DC: The National Academies Press.

"Knowing is not enough; we must apply.
Willing is not enough; we must do."

—Goethe

INSTITUTE OF MEDICINE

OF THE NATIONAL ACADEMIES

Advising the Nation. Improving Health.

THE NATIONAL ACADEMIES

Advisers to the Nation on Science, Engineering, and Medicine

The **National Academy of Sciences** is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. Upon the authority of the charter granted to it by the Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Ralph J. Cicerone is president of the National Academy of Sciences.

The **National Academy of Engineering** was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. Charles M. Vest is president of the National Academy of Engineering.

The **Institute of Medicine** was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, upon its own initiative, to identify issues of medical care, research, and education. Dr. Harvey V. Fineberg is president of the Institute of Medicine.

The **National Research Council** was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both Academies and the Institute of Medicine. Dr. Ralph J. Cicerone and Dr. Charles M. Vest are chair and vice chair, respectively, of the National Research Council.

www.national-academies.org

COMMITTEE ON AN EVIDENCE FRAMEWORK FOR OBESITY PREVENTION DECISION MAKING

SHIRIKI K. KUMANYIKA (*Chair*), Professor of Epidemiology and Associate Dean for Health Promotion and Disease Prevention, University of Pennsylvania School of Medicine, Philadelphia

DAVID B. ABRAMS, Director, Schroeder Institute for Tobacco Research and Policy Studies, American Legacy Foundation, Washington, DC; Professor, Department of Health, Behavior and Society, The Johns Hopkins Bloomberg School of Public Health, Baltimore, MD

ROSS C. BROWNSON, Professor of Epidemiology, School of Medicine (Siteman Cancer Center) and George Warren Brown School of Social Work, Washington University in Saint Louis, MO

FRANK CHALOUPKA, Professor of Economics, Director, UIC Health Policy Center, University of Illinois at Chicago

MADHABI CHATTERJI, Associate Professor of Measurement-Evaluation & Education, Teachers College, Columbia University, New York

BARBARA A. DENNISON, Director of Policy and Research Translation, Division of Chronic Disease and Injury Prevention, New York State Department of Health, Albany

CHRISTINA ECONOMOS, New Balance Chair, Childhood Nutrition, Friedman School of Nutrition Science and Policy, Tufts University, Boston, MA

STEVEN GORTMAKER, Professor of Practice of Health Sociology, Department of Society, Human Development, and Health, School of Public Health, Harvard University, Boston, MA

LAWRENCE W. GREEN, Professor, Department of Epidemiology and Biostatistics, Helen Diller Family Comprehensive Cancer Center, University of California–San Francisco

ROBERT A. HIATT, Professor and Co-Chair of Epidemiology and Biostatistics, Deputy Director, Helen Diller Family Comprehensive Cancer Center, University of California–San Francisco

WILLIAM PURCELL, III, Director, Institute of Politics, John F. Kennedy School of Government, Harvard University, Cambridge, MA

ROBERT SEGE, Professor of Pediatrics, Boston University School of Medicine, Boston, MA

HAROLD SOX, Editor Emeritus, Annals of Internal Medicine, American College of Physicians of Internal Medicine, Philadelphia, PA

ADOLFO M. VALADEZ, Assistant Commissioner, Division of Prevention and Preparedness Services, Texas Department of State Health Services, Austin

LETICIA VAN DE PUTTE, Senator, Texas State Senate, San Antonio

STEPHEN G. WEST, Professor of Quantitative and Social Psychology, Department of Psychology, Arizona State University, Tempe

Study Staff

LYNN PARKER, Scholar
LESLIE J. SIM, Program Officer
EMILY ANN MILLER, Research Associate
SAUNDRA LEE, Senior Program Assistant
MATTHEW SPEAR, Senior Program Assistant
ANTON BANDY, Financial Officer
GERALDINE KENNEDO, Administrative Assistant
LINDA D. MEYERS, Food and Nutrition Board Director

Reviewers

This report has been reviewed in draft form by individuals chosen for their diverse perspectives and technical expertise, in accordance with procedures approved by the National Research Council's Report Review Committee. The purpose of this independent review is to provide candid and critical comments that will assist the institution in making its published report as sound as possible and to ensure that the report meets institutional standards for objectivity, evidence, and responsiveness to the study charge. The review comments and draft manuscript remain confidential to protect the integrity of the deliberative process. We wish to thank the following individuals for their review of this report:

JOHN P. ALLEGRANTE, Professor and Chair, Department of Health and Behavior Studies, Columbia University, New York

ALICE S. AMMERMAN, Director, Center for Health Promotion and Disease Prevention, University of North Carolina at Chapel Hill

LISBETH BAMBERGER SCHORR, Director, Project on Effective Interventions, Harvard University, Washington, DC

KELLY D. BROWNELL, Director, Rudd Center for Food Policy and Obesity, Yale University, New Haven, CT

LINDA COLLINS, Director, The Methodology Center, Pennsylvania State University, State College

DAVID CORDRAY, Professor, Department of Psychology and Human Development, Vanderbilt University, Nashville, TN

TRISHA GREENHALGH, Professor, Primary Care and Population Sciences, University College London, UK

ABBY KING, Division Chief, Stanford Prevention Research Center, Stanford University School of Medicine, Stanford, CA

LISA M. KLESGES, Professor and Director, School of Public Health, University of Memphis

ALFREDO MORABIA, Professor of Clinical Epidemiology, Mailman School of Public Health, Columbia University, New York
LYNN SILVER, Assistant Commissioner, Chronic Disease Prevention & Control, New York City Department of Health & Mental Hygiene
JUNE STEVENS, Chair, Department of Nutrition, University of North Carolina School of Public Health, Chapel Hill
MAY C. WANG, Associate Professor, Department of Community Health Sciences, University of California–Los Angeles School of Public Health, Los Angeles

Although the reviewers listed above have provided many constructive comments and suggestions, they were not asked to endorse the conclusions or recommendations nor did they see the final draft of the report before its release. The review of this report was overseen by **ENRIQUETA C. BOND,** President Emeritus, Burroughs Wellcome Fund and **CUTBERTO GARZA,** Provost and Dean of Faculties, Boston College. Appointed by the National Research Council and the Institute of Medicine, they were responsible for making certain that an independent examination of this report was carried out in accordance with institutional procedures and that all review comments were carefully considered. Responsibility for the final content of this report rests entirely with the authoring committee and the institution.

Preface

The impetus for this study was a recognition on the part of the study's sponsors, the Institute of Medicine (IOM), and ultimately the scholars who served as committee members of the need to better align the evidence base for obesity prevention with the evidence needs of policy makers and professionals who are making decisions in this arena. Addressing the nation's obesity epidemic requires well-reasoned actions to transform the current obesity-promoting landscape such that the average adult or child is likely to avoid excess weight gain. Relevant actions are being considered and taken by decision makers in international, national, state, and local governments and by organizations and agencies in the profit and nonprofit sectors on the basis of the evidence at hand, as well as other considerations. Many of these actions involve changes in physical environments, public policies, or organizational practices. Some proposed actions will be questioned on the basis of feasibility or effectiveness; some will be strongly opposed by those who stand to lose economically or in other respects. Decision makers can use evidence to assess the potential impacts, advantages, and disadvantages of particular actions; respond to concerns; and argue their case.

The sponsors of this report have essentially called on the scientific community to rise to the challenge and transform the evidence picture to be commensurate with the needs. Limitations in the current evidence base are manifest in scientific review articles that identify a paucity of obesity prevention interventions found to be effective; in study designs that fail to address the complexity of the obesity problem when viewed from a population perspective; in the lack of data with which to cost out and prioritize various strategies; and in evidence of major policy decisions that—however justified from a practical perspective—are being taken in the absence of evidence that might be helpful and perhaps critical in determining which actions should be continued, replicated, or scaled up.

If "evidence" is given the usual interpretation of *research* evidence, the implication is that the research community—those expected to be the keepers, judges, and providers of evidence—need to do something different so the research will have more practical utility. Several questions then arise: What is that "something different"? What is missing from the evidence that is available now? How can and should identified gaps be filled? And ultimately, can a more useful approach to evidence be described in a way that (1) engenders agreement and respect within the scientific community, (2) has practical utility, and (3) therefore, also reassures decision makers of the validity of the approach and motivates its use? These are the fundamental issues addressed in this report, and the committee has wrestled long and hard to achieve a result that will indeed transform the way obesity prevention research is defined and undertaken to better meet potential user needs.

The committee deliberated extensively about what actually constitutes good research with respect to complex population problems such as obesity, and when and how currently held assumptions about research quality require rethinking to make research results usable in the real world. We considered how concepts of evidence can be broadened to encompass information that is sound but not necessarily what comes to mind when one thinks of evidence in traditional terms. We examined why the problems we identified with the evidence framework for obesity prevention exist and what might stand in the way of implementing a new framework. We did not find shortcuts. From a practical perspective, aligning evidence with user needs and filling evidence gaps will require dedicated effort, time, money, and infrastructure, as outlined in this report's recommendations—inputs that will have broader relevance and utility when also applied to other population health problems. We did identify new approaches, some reflected in efforts already under way and others that could be implemented immediately. These new approaches include talking and sharing across disciplines; thinking about relevance and the connections among the myriad population- and community-level influences that can be focal points of obesity prevention initiatives; and taking advantage of the opportunity to evaluate ongoing initiatives, that is, to obtain the practice-based evidence that is intrinsically relevant to natural settings.

The committee was a truly tireless group with a shared commitment to bringing the full potential of scientific excellence, broadly defined, to bear on addressing populationwide obesity, including the application of learning from exemplary efforts of this type in other areas of population health. The committee was aided by a stellar roster of colleagues who gave us the benefit of their thinking and experience at two public workshops: Stephanie Coursey Bailey, Allan Best, Steven Daniels, Ana Diez-Roux, Joseph Hagan, Jr., Jack Homer, Susan Jebb, David L. Katz, Bobby Milstein, J. Michael Oakes, Joshua Sharfstein, Lynn Silver, and June Stevens. As the person given the privilege of chairing this effort, I express my sincerest gratitude to all of these colleagues, as well as to the study's sponsors—Kaiser Permanente, the Robert Wood Johnson Foundation, and the Centers for Disease Control and Prevention—for the vision that was inherent in the committee's statement of task. In addition, and on behalf of the committee as a whole, I commend and thank our very capable IOM staff who applied their patience, wisdom, and hands-on assistance liberally and expertly throughout the course of this effort: Lynn Parker, Study Director; Leslie Sim, Program Officer; Emily

Ann Miller, Research Associate; Matthew Spear and Saundra Lee, Senior Program Assistants; and Linda D. Meyers, Food and Nutrition Board Director.

The results of the committee's efforts are offered to decision makers and to the research community for use, scrutiny, and debate, recognizing that any effort of this type will be a work in progress that will benefit from interactions with the people and contexts to whom and to which it is relevant. In keeping with the spirit of this report, the debate should be far from "academic," deriving its energy and arguments by remaining close to the action in the field.

Shiriki K. Kumanyika, *Chair*
Committee on an Evidence Framework for
Obesity Prevention Decision Making

Contents

Summary

In the United States, data from the 2007-2008 National Health and Nutrition Examination Survey indicate that nearly 34 percent of adults aged 20 years or older are obese and about 68 percent are overweight or obese. Among children and adolescents aged 2 to 19, nearly 17 percent are obese and close to 32 percent are overweight or obese. The epidemic reflected by these figures poses major challenges for policy makers, public health professionals, and other decision makers.

Federal health authorities have called for decisive action to address this epidemic for more than a decade, and their appeals have increased in urgency in recent years. Their alarm is linked to the adverse health outcomes of obesity and overweight, including potential negative effects on quality of life and life expectancy. It is also triggered by the negative economic impact of these levels of obesity. A recent analysis calculated that the annual national medical burden of obesity is nearly 10 percent of all medical spending and could have risen to $147 billion per year by 2008.

INTRODUCTION

Policy makers and practitioners who take action to address this epidemic need evidence to explain why the actions are needed and to help determine what actions to take and how to go about them. They also need evidence to determine whether the intended results are obtained, who benefits, and whether there are positive or negative side effects that deserve attention. Ideally, there would be ready sources of such evidence at hand relevant to these information needs.

Unfortunately, a review of the existing evidence base reveals a striking contrast between the high prevalence and consequent importance of addressing obesity and the paucity of the knowledge base with which to inform prevention efforts. Specifically for evidence-based obesity prevention efforts, a body of intervention research on policy and environmental approaches is largely absent from the literature.

The types of questions that are being asked by decision makers interested in obesity prevention cannot be answered without broadening the concept of evidence to include new ways of using traditional or existing information; information drawn from a wider range of sources; and newly generated evidence that is more focused on the design, implementation, and outcomes of policies and programs to prevent obesity—whether initiated in research or practice settings.

THE NEED FOR A NEW FRAMEWORK

The scenarios in Boxes S-1 and S-2, respectively, illustrate the need to assist decision makers in finding evidence to inform decisions and to help researchers determine useful areas of study. Considerable progress has been made in building on the strengths of methods used in evidence-based medicine to develop research methods and evidence standards applicable to decision making about other population-level health problems. Leveraging these accomplishments holds the potential to advance understanding of how to develop and evaluate evidence for use in decision making on obesity prevention.

The Institute of Medicine Committee on an Evidence Framework for Obesity Prevention Decision Making was formed to address these challenges. The overall charge of the committee was to develop a framework for evidence-informed decision making in obesity prevention, with a focus on assessing policy, environmental, and

Box S-1
The Need for Evidence to Support Decision Making on Obesity Prevention

Imagine that you are the mayor, or a health commissioner, in a city in which both children and adults have high rates of obesity—perhaps higher than in some peer cities that appear to have a better handle on the problem. You need to decide which of a spectrum of actions to take and how to justify these actions against some inevitable opposition from various stakeholders. Consumer advocates are calling for policies that require posting of calories on the menu boards of fast food restaurants. Some city council members are calling for taxes on soft drinks to lower consumption levels but also to raise revenue to offset budget deficits. The school board is debating whether to make the collection of child weight-for-height data mandatory and to send the information home to parents.

Ideally, among the many considerations factoring in to your decision making, you or your staff would readily find or have at hand a rich and frequently replenished set of reports to support your decision about a particular strategy or a choice among options. These reports would relate to the types of policy and environmental changes you are considering and provide an idea of their pros and cons—for example, estimated benefits, unintended adverse consequences, costs, and the practicalities of implementation. You would find few such reports, however.

Box S-2
The L.E.A.D. Framework and New Opportunities for Research

Imagine that you are a researcher who has become aware of the value of conducting research with an impact on public health. You have seen a call for proposals for policy-oriented research related to obesity or have recognized the need to expand your current obesity research to incorporate more of a multilevel perspective. Your research might focus on obesity or some other health issue or be in another field, such as city planning, education, or law. What are innovative ways to think about relevant research questions that would be fundable and publishable? Are there ways to research complex, big-picture questions that need answering? How far can research that might be relevant to those questions deviate from the status quo with respect to methods considered to be the gold standard in your field? How can you ensure that your research is responsibly designed to assess a policy adequately? One purpose of the L.E.A.D. framework is to stimulate new ways of thinking about research that can yield answers to such questions.

community interventions that influence food, eating, and physical activity. Inherent in this charge was a recognition that, while treatment and prevention focused on the individual remain relevant, there is a growing need for obesity prevention strategies that focus on whole populations—multicomponent, multilevel strategies that can favorably impact communities or other complex systems.

In developing the L.E.A.D. framework (for **L**ocate Evidence, **E**valuate Evidence, **A**ssemble Evidence, Inform **D**ecisions), the committee was instructed to provide an overview of the current nature of the evidence base; identify the challenges faced in integrating scientific evidence into the broader array of factors that influence community interventions and policy change; provide practical, action-oriented recommendations for using this framework to choose, implement, and evaluate obesity efforts; identify new research and evaluation tools and methods, and existing ones that can be deployed more effectively; and develop a plan for communicating, disseminating, evaluating, and refining the framework. The committee's charge emphasized the need for a framework that guides decision making on children and adults. The committee also was directed to focus on the role of a systems perspective in making obesity prevention decisions, as well as to contribute to more general efforts to address complex, multifactorial public health challenges.

STUDY APPROACH

In responding to its charge, the committee's main goal was to support decision makers in choosing and implementing obesity prevention interventions or in assessing the outcomes of interventions already in place or under way. This report has two primary audiences: (1) decision makers and the intermediaries who assist them in making decisions and (2) those who conduct research relevant to obesity prevention or who evaluate existing programs. Other important audiences include research funders and publishers of research findings.

Two caveats are necessary as a prelude to presenting the L.E.A.D. framework. First, this framework does not offer shortcuts that can provide high-quality evidence as rapidly as it might be needed. The framework offers suggestions for ways to increase flexibility and broaden perspectives while adhering closely to concepts of what makes evidence credible as well as useful. Second, this is not a framework that can be used as an algorithm for making decisions; rather, it is a framework for using evidence to inform decision making. A richer, more relevant evidence base will inform and vastly improve decision making, but will not make the design, selection, and assessment of interventions automatic or free decision makers from exercising judgment.

This report's audiences need answers to two fundamental questions:

- How can evidence that is currently available and potentially relevant to decisions on obesity prevention be identified, evaluated, and compiled in ways that will best inform decision makers?
- How can more evidence be developed that is of high quality and framed to be directly relevant to decision making on obesity prevention?

USING THE L.E.A.D. FRAMEWORK

The L.E.A.D. framework is illustrated in Figure S-1. This framework includes concepts and approaches that are standard procedure in the development of practice guidelines. However, the committee incorporated major innovations that expand and enhance these approaches for use in policy and programmatic decision making on complex public health problems like obesity. The framework and its supporting narrative:

- explain why it is critical to use a systems perspective;
- characterize the types of questions policy makers ask;
- broaden the concept of evidence;
- reframe the definition of quality of evidence to accord with the type of evidence;
- recommend ways to consider other relevant information when evidence is limited;
- propose a template for assembling evidence; and
- highlight opportunities to generate new and relevant evidence.

The framework leads decision makers and researchers through a series of steps, from specifying questions to informing decisions. However, a user can begin at any point on the framework and return to earlier steps, for example, moving back from the "assemble" step to elaborate further on the questions and broaden the search for evidence.

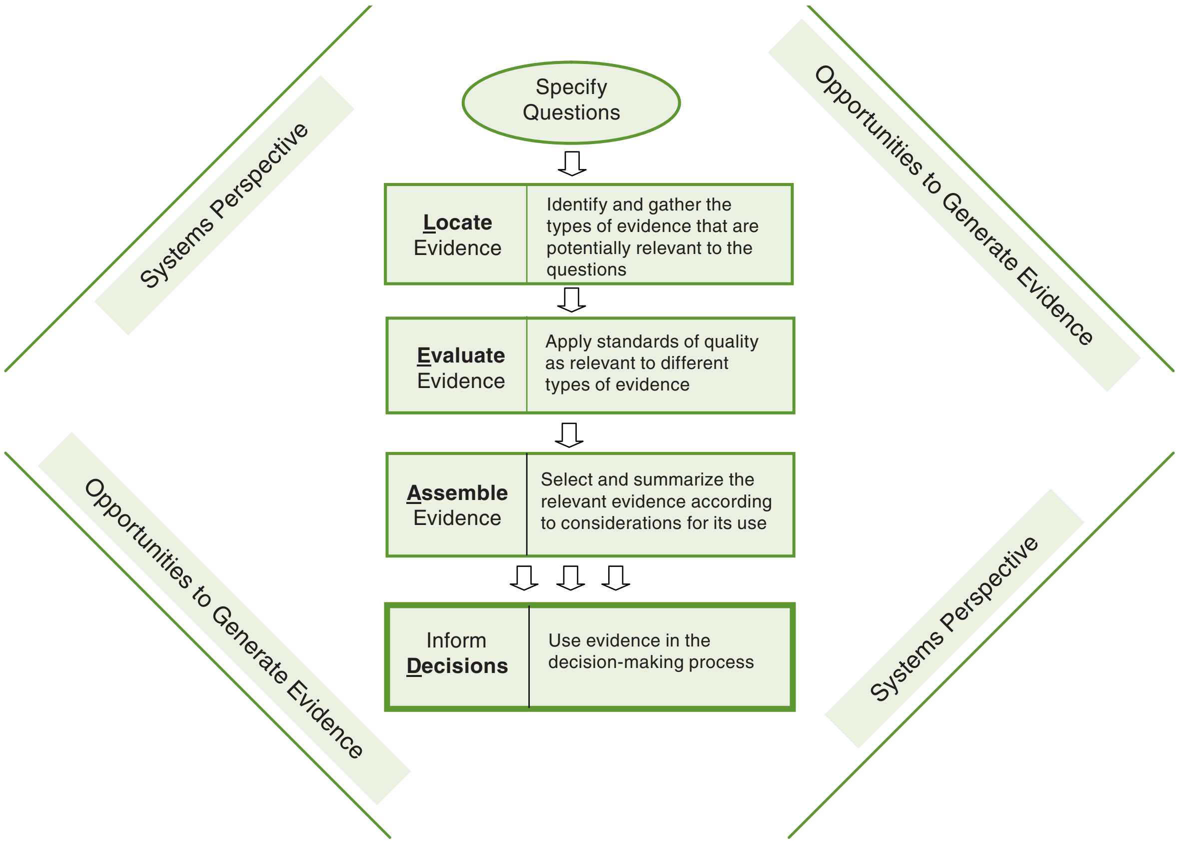


FIGURE S-1 The L.E.A.D. framework.

The presence of the term "systems perspective" surrounding the steps in the L.E.A.D. framework in Figure S-1 emphasizes the importance of taking such a perspective throughout the process of applying the framework—from specifying questions to guide the search for evidence to taking advantage of opportunities to generate new evidence. The presence of the term "opportunities to generate evidence" surrounding the framework steps emphasizes the importance of identifying research opportunities throughout the process. Following are brief explanations of each aspect of the L.E.A.D. framework in Figure S-1. The "Key Action" boxes are primarily for the use of decision makers and their intermediaries (i.e., individuals who may assist decision makers in using the framework).

Systems Perspective

Key Action: Approach every aspect of decision making and research with a fuller appreciation of the complex context in which programs and policies are implemented.

The real world is a complex system, and the many influences on the energy-balance equation are all interacting simultaneously. A systems perspective helps decision makers and researchers think broadly about the whole picture rather than merely studying the component parts in isolation. Such a perspective can help to frame, explain, and resolve complex problems such as obesity. It can lead to a better understanding of interactions and highlights the importance of taking into account the context in which public health problems occur and how that context may affect the implementation and impact of interventions.

A systems perspective can enhance the ability to develop and use evidence effectively and suggest actions with the potential to effect change. It can allow the forecasting of potential consequences of not taking action, possible unintended effects of interventions, the likely magnitude of the effect of one or more interventions, conflicts between or complementarity of interventions, and priorities among interventions. Systems maps—in which relevant factors and pathways are illustrated graphically—can help organize an overwhelming array of potential influences and pathways relevant to the determinants of obesity.

Specifying Questions

Key Action: Identify the questions that need to be answered in order to make the needed decision.

There are many kinds of decisions that are likely to be made on obesity prevention at the local, state, and national levels. The common applications of evidence for which the L.E.A.D. framework can be useful are:

- justifying interventions;
- setting priorities with respect to specific outcomes;
- justifying particular actions in high-risk population segments;
- distinguishing actions that are likely to be effective from those that are not;
- quantifying likely impacts;
- estimating costs and cost-effectiveness;
- anticipating unintended consequences; and
- understanding implementation variables, i.e., what to do and how to do it.

Thus, decisions about the selection, implementation, and evaluation of interventions are closely tied to questions about why an intervention is needed and how a specific intervention can be expected to work in a given context. The L.E.A.D. framework adapts an evidence-based public health typology to differentiate among three categories of interrelated questions—"Why," "What," and "How." These general questions can help decision makers identify related questions that may inform their policy and program decisions, underscore the fact that no one type of evidence will be able to answer all relevant questions, and facilitate the process of locating useful evidence for different questions.

"Why" questions relate to the overall question of "Why should we do something about this problem in our situation?" "What" questions ask, "What specifically should we do about this problem?" "How" questions call for examining "How do we implement this information for our situation?"

Locating Evidence

Key Actions: (1) Based on the questions that need to be answered, determine all the types of evidence that could be useful in answering them. (2) Think broadly about the sources of these types of evidence, including the potential for obtaining relevant evidence from other disciplines.

Locating evidence requires a clear concept of the types of information that may be useful for a particular purpose, as well as an awareness of where the information can be found. The framework calls for expanding the evidence paradigm by broadening the perspective on forms of evidence that are potentially relevant and useful, taking full advantage of available research methods for studying population problems, not just those used in medical research. Some forms of available evidence may be underutilized by those who conduct research on obesity prevention because they are unfamiliar to researchers in the biomedical or public health fields.

Locating evidence also requires awareness of and access to appropriate information resources. Databases used by public health researchers and practitioners typically incorporate data from numerous disciplines but may still miss many potentially useful sources. For example, compilations from economics, education, business, and law and information from newspapers, government documents, and reports from community agencies and programs may not be obvious sources of evidence for obesity prevention researchers.

Evaluating Evidence

Key Actions: (1) Recognize the importance of evaluating the quality of the evidence gathered to answer the specified questions. (2) When evaluating evidence, use criteria that are appropriate and established for assessing the quality of that particular type of evidence. (3) In evidence evaluation, pay attention to both the level of certainty (internal validity) and generalizability (external validity) of the evidence.

In the L.E.A.D. framework, the key objectives in evaluating evidence are determining the level of certainty of the causal relationship between an intervention and the observed outcomes (or internal validity) and generalizability to other individuals, settings, contexts, and time frames (or external validity). The level of certainty needed will vary depending on the question. For example, some low-cost interventions with minimal potential for harm may require less certainty than those that are costly or carry the risk of serious harm.

A key issue in evaluating evidence is aligning the question(s) of interest with the appropriate outcome(s). Outcomes may be short-term, intermediate, or long-term. For example, some obesity prevention decisions will require evidence that relates directly to weight outcomes, while others may allow evidence related to intermediate behavioral outcomes.

Evaluations of interventions should also be sensitive to the nature of the intervention. For example, outcomes that are farther downstream from a policy change might also be of interest but would be less reflective of the specific effect of the change given the other influences that might have intervened in the interim. Quality considerations in assessing evidence, while based on the same principles, vary with the form or source of the evidence. Different types of evidence require different approaches to judging validity and other aspects of quality. Evaluating relevance to the context to which the question applies may be an additional critical step in any assessment of evidence for complex population-level interventions, requiring the application of criteria to judge the generalizability of the evidence.

Assembling Evidence and Informing Decisions

Key Actions: (1) Develop a transparent and comprehensive summary of the evidence available on the decision that must be made, based on the information gathered by following the L.E.A.D. framework. (2) Include in this summary the question(s) asked by the decision maker; the strategy for gathering and selecting the evidence; an evidence table showing the sources, types, and quality of the evidence and the outcomes reported; and a concise summary of the evidence on why an action should be taken, what that action should be, and how it should be taken. (3) If obesity prevention actions must be taken when evidence is limited, examine the potential for blending the limited evidence with theory, professional experience, and local wisdom. (4) Use this summary to inform the decision-making process.

Once the potentially relevant evidence has been located and evaluated with the more broadly based, interdisciplinary view called for by the L.E.A.D. framework, the evidence needs to be synthesized and summarized to help inform the decision based on the question(s) asked. Taken together, the results of the overall evaluation of the evidence should assist in providing answers to why action should be taken, what action to take, and how to take it. Implicit in this process is the understanding that more than one type and source of evidence will be needed to inform decision making, and that there will be inevitable trade-offs between level of certainty and generalizability.

A standardized approach to assembling the results of evidence gathering and evaluation facilitates transdisciplinary discussion among stakeholders and presents evidence to decision makers in a usable form. A uniform language for drawing and describing conclusions signals the use of a uniform set of procedures to evaluate the evidence and improves clarity in communication. Clarity is particularly important when expertise from many different disciplines is required, each with its own jargon and methods.

The intent of the transparent and uniform reporting of the process and conclusions of evidence gathering and synthesis is to provide decision makers with information they can understand as they make choices among alternative policies and programs. The proposed report template aligns with the main elements of the framework, describing the question(s) asked by the decision maker, the strategy for gathering and selecting evidence, and the evaluation of the evidence, and ends with a summary of the synthesized evidence. This summary should address the broad categories of effectiveness in the proposed setting, population affected and potential impact, and implementation.

The ability to use evidence to inform decision making depends in part on the availability of relevant evidence. Use of the L.E.A.D. framework can broaden what is considered to be useful, high-quality evidence and gradually increase the amount of such evidence. Yet, despite the best efforts to amass available evidence, those grappling with an emerging problem such as obesity will face decisions that must be made on the basis of inconsistent or incomplete evidence. Similar to what has been observed in tobacco control, one can anticipate cycles of planning that begin with incomplete evidence, blended with theory, expert opinion, experience, and understanding of local traditions and the probable response to proposed actions, and extend to evaluating the consequences of interventions. Decision making is complex and takes many factors, in addition to evidence, into account. The goal is to enable the best possible use of evidence within decision making processes.

Opportunities to Generate Evidence

Key Actions: (1) Take full advantage of opportunities to generate evidence from ongoing policy and practice. (2) If obesity prevention actions are taken when the evidence is very limited, evaluate the success of the intervention and build credible evidence for use in future decision making. (3) Treat natural experiments, emerging innovations, and ongoing programs as potential sources of useful evidence. (4) Consider forms of evidence and research designs from a variety of disciplines, including systems approaches that can handle complexity. (5) Explore research designs that can be used as alternatives to randomized experiments and that may be more feasible in relation to complex environmental and policy interventions. (6) When reporting results of obesity prevention efforts, include useful aspects of the research related to its generalizability to individuals, settings, contexts, and time frames.

Researchers in a variety of fields, as well as those who fund and publish their research, are among the intended users of the L.E.A.D. framework. In addition to fostering a systems approach and making the best possible use of diverse types of relevant evidence, application of the L.E.A.D. framework can suggest opportunities for research. This may occur during any step of the process.

What is usually regarded as the current evidence base for obesity prevention is limited in both size and utility. New approaches based on the expanded perspective outlined in this report are urgently needed to address these limitations. Among other problems, the limited nature of evidence on effectiveness may be interpreted as a lack of effectiveness, when what it actually indicates is that the degree of effectiveness is not yet known.

Research to inform answers to the questions that drive the use of the L.E.A.D. framework should take full advantage of opportunities to generate evidence from ongoing policy and practice, as well as evidence and research designs from a variety of disciplines, including systems approaches that can handle complexity. This report includes an extensive discussion of research designs that can be used as alternatives to randomized experiments and that may be more feasible and applicable for studying complex environmental and policy interventions such as those aimed at obesity prevention. The report also contains a discussion of the need to take advantage of natural experiments and emerging and ongoing interventions as sources of practice-based evidence. Finally, research related to obesity prevention and related public health issues is inherently transdisciplinary.

NEXT STEPS

An effective communication and dissemination plan for the L.E.A.D. framework should focus on three major objectives: (1) introducing the purpose of the framework to decision makers, their intermediaries, researchers, and others; (2) raising awareness of the importance of marshaling evidence within both policy and practice environments and research environments to support decision making for complex public health challenges like obesity; and (3) creating support for integrating the framework into current policy making and research practice. The knowledge gained as the framework is disseminated, implemented, and evaluated can be used to refine it and its application. A strategic evaluation and refinement plan for the framework can measure the adoption and utilization of the framework by the intended users and evaluate its impact on research and decision-making strategies. A well-executed communication, dissemination, evaluation, and refinement plan for the L.E.A.D. framework will help in adapting it to real-world policy making and practice.

RECOMMENDATIONS

The United States has made progress toward translating science to practice in the brief time since the obesity epidemic was officially recognized. But the pace of this translation has been slow relative to the scope and urgency of the problem and the associated harms and costs. With this in mind, the committee makes the following recommendations for assisting decision makers and researchers in using the current evidence base for obesity prevention and for generating more, and more useful, evidence.

Utilize the L.E.A.D. Framework

Recommendation 1: Decision makers and those involved in generating evidence, including researchers, research funders, and publishers of research, should apply the L.E.A.D. framework as a guide in their utilization and generation of evidence to support decision making for complex, multifactorial public health challenges, including obesity prevention.

Key assumptions that should guide the use of the framework include the following:

- A systems perspective can help in framing and explaining complex issues.
- The types of evidence that should be gathered to inform decision making are based on the nature of the questions being asked, including Why? ("Why should we do something about this problem in our situation?"), What? ("What specifically should we do about this problem?"), and How? ("How do we implement this information for our situation?"). A focus on subsets of these questions as a starting point in gathering evidence explicitly expands the evidence base that is typically identified and gathered.
- The quality of the evidence should be judged according to established criteria for that type of evidence.
- Both the level of certainty of the causal relationship between an intervention and the observed outcomes and the intervention's generalizability to other individuals, settings, contexts, and time frames should be given explicit attention.
- The analysis of the evidence to be used in making a decision should be summarized and communicated in a systematic, transparent, and transdisciplinary manner that uses uniform language and structure. The report on this analysis should include a summary of the questions asked by the decision maker; the strategy for gathering and selecting the evidence; an evidence table showing the sources, types, and quality of the evidence and the outcomes reported; and a concise summary of the synthesis of selected evidence on why an action should be taken, what that action should be, and how it should be taken.
- If action must be taken when evidence is limited, this incomplete evidence can be blended carefully and transparently with theory, expert opinion, and collaboration based on professional experience and local wisdom to support making the best decision.

Sustained commitments will be needed from both the public and private sectors to achieve successful utilization of the various elements of the L.E.A.D. framework in future evidence-informed decision making and evidence generation. This responsibility lies with the academic and research community, as well as with government and private funders and the leadership of journals that publish research in this area. Necessary supports will include increasing understanding of systems thinking and incorporating it into research-related activities, creating and maintaining resources to support the utilization of evidence, establishing standards of quality for different types of evidence, and supporting the generation of evidence, each of which is described in more detail below. Finally, it will be necessary to communicate, disseminate, evaluate, and refine the L.E.A.D. framework.

Incorporate Systems Thinking

Recommendation 2: Researchers, government and private funders, educators, and journal editors should incorporate systems thinking into their research-related activities.

To implement this recommendation:

- Researchers should use systems thinking to guide the development of environmental and policy interventions and study designs.
- Government and private funders should encourage the use of systems thinking in their requests for proposals and include systems considerations in proposal evaluations.
- Universities, government agencies such as the U.S. Centers for Disease Control and Prevention, and public health organizations responsible for educating public health practitioners and related researchers should establish training capacity for the science and understanding of systems thinking and the use of systems mapping and other quantitative or qualitative systems analysis tools.
- Journal editors should encourage the use of systems thinking for addressing complex problems by developing panels of peer reviewers with expertise in this area and charging them with making recommendations for how authors could use systems thinking more effectively in their manuscripts.

Build a Resource Base

Recommendation 3: Government, foundations, professional organizations, and research institutions should build a system of resources (people, compendiums of knowledge, registries of implementation experience) to support evidence-based public policy decision making and research for complex health challenges, including obesity prevention.

To implement this recommendation:

- The Secretary of Health and Human Services, in collaboration with other public- and private-sector partners, should establish a sustainable registry of reports on evidence for environmental and policy actions for obesity prevention. Integral to this registry should be the expanded view of evidence for decision making on obesity prevention proposed in this report and the sharing of experiences and innovative programs as the evidence evolves. A service provided by this registry should be periodic synthesis reviews based on mixed qualitative and quantitative methods.

- The Secretary of Health and Human Services, in collaboration with other public- and private-sector partners, should develop and fund a resource for compiling and linking existing databases that may contain useful evidence for obesity prevention and related public health initiatives. This resource should include links to data and research from disciplines and sectors outside of obesity prevention and public health and to data from nonacademic sources that are of interest to decision makers.

Establish Standards for Evidence Quality

Recommendation 4: Government, foundations, professional organizations, and research institutions should catalyze and support the establishment of guidance on standards for evaluating the quality of evidence for which such standards are lacking.

To implement this recommendation:

- Government and private funders should give priority to funding for the development of guidance on standards for evaluating the quality of the full range of evidence types discussed in this report that are useful in making obesity prevention decisions, especially those for which the scientific literature is limited.
- Professional organizations and research institutions should encourage and bring attention to efforts by faculty, researchers, and students to establish guidance in this area.

Support the Generation of Evidence

Recommendation 5: Obesity prevention research funders, researchers, and publishers should consider, wherever appropriate, the inclusion in research studies of a focus on the generalizability of the findings and related implementation issues at every stage, from conception through publication.

To implement this recommendation:

- Those funding research in obesity prevention should give priority to support for studies that include an assessment of the limitations, potential utility, and applicability of the research beyond the particular population, setting, and circumstances in which the studies are conducted, including by initiating requests for applications and similar calls for proposals aimed at such studies. Additional ways in which this recommendation could be implemented include adding criteria related to generalizability to proposal review procedures and training reviewers to evaluate generalizability.

- Obesity prevention researchers and program evaluators should give special consideration to study designs that maximize evidence on generalizability.
- Journal editors should provide guidelines and space for authors to give richer descriptions of interventions and the conditions under which they are tested to clarify their generalizability.

Recommendation 6: Research funders should increase opportunities for those carrying out obesity prevention initiatives to measure and share their outcomes so others can learn from their experience.

To implement this recommendation:

- Organizations funding or sponsoring obesity prevention initiatives—including national, regional, statewide, or local programs; policy changes; and environmental initiatives—should provide resources for obtaining practice-based evidence from innovative and ongoing programs and policies in a more routine, timely, and systematic manner to capture their processes, implementation, and outcomes. These funders should also encourage and support assessments of the potential for evaluating the most innovative programs in their jurisdictions and sponsor scientific evaluations where the opportunities to advance generalizable evidence are greatest.
- Research funders, researchers, and journal editors should assign higher priority to studies that test obesity prevention interventions in real-world settings in which major contextual variables are identified and their influence is evaluated.

Recommendation 7: Research funders should encourage collaboration among researchers in a variety of disciplines so as to utilize a full range of research designs that may be feasible and appropriate for evaluating obesity prevention and related public health initiatives.

To implement this recommendation:

- As part of their requests for proposals on obesity prevention research, funders should give priority to and reward transdisciplinary collaborations that include the creative use of research designs that have not been extensively used in prevention research but hold promise for expanding the evidence base on potential environmental and policy solutions.

Communicate, Disseminate, Evaluate, and Refine the L.E.A.D. Framework

Recommendation 8: A public–private consortium should bring together researchers, research funders, publishers of research, decision makers, and other stakeholders to discuss the practical uses of the L.E.A.D. framework, and develop plans and a timeline for focused experimentation with the framework and for its evaluation and potential refinement.

To implement this recommendation:

- Interested funders should bring together a consortium of representatives of key stakeholders (including decision makers, government funders, private funders, academic institutions, professional organizations, researchers, and journal editors) who are committed to optimizing the use of the current obesity prevention evidence base and developing a broader and deeper base of evidence.
- This consortium should develop an action-oriented plan for funding and implementing broad communication, focused experimentation, evaluation, and refinement of the L.E.A.D. framework. This plan should be based on the major purposes of the framework: to significantly improve the evidence base for obesity prevention decision making on policy and environmental solutions, and to assist decision makers in using the evidence base.

1

Introduction

KEY MESSAGES

- There is increasing recognition that overweight and obesity are not just problems of individuals, but also societywide problems that have serious health consequences and costs and affect some population groups more than others. Acting on this recognition will require multifaceted, population-based changes in the social and environmental[1] factors that influence energy intake and expenditure.
- There is a pressing need to act on the problem of obesity, but there is also a large gap between the types and amount of evidence deemed ideal as the basis for taking action and the types and amount of information available to meet that need.
- The Committee on an Evidence Framework for Obesity Prevention Decision Making was formed to address the issues of how to identify, use, and evaluate evidence to inform obesity prevention decisions and how to generate evidence where it is lacking.

This report focuses on the challenge of obesity prevention, viewed from a population or public health perspective. It is about how to increase the likelihood that the decisions made in the many domains that influence obesity prevention will bring about the desired results. Observations of increases in weight among adults and youth globally within just a few decades have forced a paradigm shift in which overweight and obesity are understood not just as problems of individuals, but also as societal problems, affecting whole populations. Causal explanations and pathways to potential solutions have changed accordingly. The relatively rapid upward trends in average weight levels can in large part be explained by changes in social and environmental

[1] The term "environmental" refers in this report to people's cultural, socioeconomic, and physical contexts.

variables that influence energy intake and expenditure—and the evidence of these changes is readily available and consistent.

Changes in the food and physical activity environments have caused sustained positive energy balance, leading to weight gain. For example, energy-dense foods and beverages have become more inexpensive, heavily advertised, and readily available than in the past, and their portion sizes continue to increase. Technological innovation in the home, workplace, and schools has reduced the need for physical effort, leading to sedentary behavior. Many schools have reduced or eliminated recess and physical education programs (IOM, 2005). The field of obesity prevention has made advances in determining the causes of the epidemic, but the complex interrelationships among those many causes have been difficult to address adequately within a given intervention. For this reason, a systems approach to obesity prevention is advocated in this report (see Chapter 4).

Taking a population perspective on obesity does not negate the importance of genetic or biologically mediated influences on the problem. But the increasing prevalence of obesity in diverse populations reveals the limitations of biological controls on energy balance in the face of high caloric availability and diminished caloric expenditure through routine physical activity. Nor does the recognition of environmental influences on obesity negate the importance of individual weight control efforts; however, the level of effort required makes individuals' attempts to control their weight notoriously ineffective over the long term. Changes in aspects of society that encourage excess caloric consumption and sedentary behavior are therefore necessary to support individual efforts to achieve and maintain energy balance. In fact, measures that have been recommended for populationwide obesity prevention would enhance the potential for obesity treatment to be effective over the long term. Moreover, unless both weight gain trends and average population weight levels decline, progress achieved through treatment for those who are already obese will be offset by the tide of new people needing such treatment.

Calls for multifaceted efforts to prevent obesity have been made repeatedly by governments, scientific and professional societies, advocacy groups, and funding agencies at the local, state, national, and international levels. The case for obesity prevention is especially strong in relation to children and youth. The case was made for U.S. children in a 2005 Institute of Medicine (IOM) report *Preventing Childhood Obesity: Health in the Balance*—a report that was generated in response to a congressional request for a national action plan addressing the problem. Protecting children from harm and creating social and environmental conditions that support healthy growth and development are fundamental responsibilities of all societies. The physical and psychosocial consequences of obesity begin in childhood and may have a lifelong impact, especially—although not only—if obesity persists into adulthood. Indeed, most obesity and obesity-related diseases develop during adulthood—an important justification for obesity prevention directed at the adult population as well. Another critical justification is that an obese adult population is an incubator, biologically and environmentally, for childhood obesity.

This chapter describes the charge and scope of the IOM Committee on an Evidence Framework for Obesity Prevention Decision Making, as well as the approach the committee took to fulfill its task. It also sets forth the case for viewing obesity as a problem of whole populations and for conceptualizing prevention strategies from a population perspective. Finally, it provides an orientation to the broad array of decisions of potential interest with respect to obesity prevention.

STUDY CHARGE, APPROACH, AND SCOPE

Several prior IOM committees have faced the need to work around evidence gaps and the fact that evidence hierarchies applied to medical treatment have limitations for assessing population-based preventive interventions. A committee tasked to develop a national action plan for child obesity prevention ultimately developed a set of principles to guide its process for reviewing evidence and making recommendations. These principles underscored the need for evidence to "inform and guide policy and programmatic decisions." That committee also emphasized that the "absence of experimental evidence does not indicate a lack of causation or the ineffectiveness of an obesity prevention intervention," that circumstances warrant taking preventive actions "even if there is as-yet-incomplete scientific evidence on the interventions," but also that "there is an obligation to accumulate appropriate evidence not only to justify a course of action but to assess whether it has made a difference" (IOM, 2005, pp. 111, 114). The committee defined evidence broadly to include several types of quantitative and qualitative research studies or data that it consulted as appropriate to the type of question being answered and used an integrated approach that included multiple types of evidence, drawing on an approach then under development by an international organization. One of the hallmarks of the committee's report was the use of historical case analyses—a social science method—applied to other complex public health scenarios as an alternative to typical traditional biomedical research studies.

The specific need for evidence on the effectiveness of interventions undertaken in various sectors and contexts was addressed by the IOM Committee to Assess Progress in Preventing Childhood Obesity (IOM, 2007). The IOM Committee on Food Marketing to Children and Youth (IOM, 2006) also highlighted several evidence issues and gaps that influenced its systematic review of evidence linking food and beverage marketing to childhood obesity or its behavioral determinants. Most recently, the IOM Committee on Local Government Actions to Prevent Childhood Obesity (IOM, 2009) followed the previous IOM recommendation to use the "best *available* evidence" rather than wait for the "best *possible* evidence" (IOM, 2005, p. 3) to recommend obesity prevention approaches based on potential effectiveness, impact, and suitability for local settings and authority. In light of the limitations of the current evidence base, that committee chose not to rank the recommended approaches. This decision was based on a desire to avoid the risk of prematurely devaluing approaches that are promising but still evolving and recognition that many context-specific factors, including the level of community support and the costs involved, determine which approaches are the best fit for a given community.

These evidence issues are not new and have already been the focus of many efforts in the field of public health in relation to other complex health problems. However, the issues are far from resolved and feel very new to the field of obesity because it has heretofore been oriented to treatment rather than public health and prevention. Considering these issues in relation to obesity prevention has the potential to advance the field of public health generally while also meeting the immediate need for clarity on evidence issues related to addressing the obesity epidemic.

Study Charge

Recognizing the importance of the issues outlined above, the IOM formed the Committee on an Evidence Framework for Obesity Prevention Decision Making, with funding from the Centers for Disease Control and Prevention (CDC), Kaiser Permanente, and the Robert Wood Johnson Foundation. The IOM appointed the 16-member committee in November 2008 to address the issues of how to identify, use, and evaluate evidence to inform obesity prevention decisions and how to generate evidence where it is lacking. These issues were identified as being of major concern or relevance to a range of potential users: decision makers in government and the private sector, practitioners in public health and other sectors, and program evaluators and other public health researchers, as well as research funders, journal editors, and the media, which determine, respectively, what research is funded, published, and disseminated. The committee members were selected to include academic researchers, practitioners, and policy makers, representing content expertise in various aspects of child and adult obesity prevention interventions; public health policy practice and research more generally; and research methodology from the perspectives of biomedicine, education, and psychology.

The committee was tasked to:

- provide an overview of the nature of the evidence base for obesity prevention as it is currently construed;
- identify the challenges associated with integrating scientific evidence with broader influences on policy and programmatic considerations;
- provide a practical and action-oriented framework of recommendations for how to select, implement, and evaluate obesity prevention efforts;
- identify ways in which existing or new tools and methods can be used to build a useful and timely evidence base appropriate to the challenges presented by the obesity epidemic, and describe ongoing attempts to meet these challenges;

- develop a plan for communicating and disseminating the proposed framework and its recommendations; and
- specify a plan for evaluating and refining the proposed framework in current decision-making processes.

Study Approach

A variety of sources informed the committee's work. The committee's deliberations drew on the findings and conclusions of the above-referenced IOM reports, and on other resources from within and outside the United States related to evidence-based public health and public policy as identified through extensive searches of the literature and a commissioned paper. Two public workshops were convened, in January and March 2009, to obtain additional input on evidence uses and challenges and on systems thinking and methodology (see Appendix F for the workshop agendas).

Obesity prevention differs from obesity treatment in the types of questions being addressed and answers being sought, the contexts in which the questions are asked and the answers are applied, and the types of research and information that may be relevant. The realities of using evidence to inform decisions about obesity treatment and prevention also differ in the availability of relevant research. In addition, because many influences on obesity are outside the health domain, information relevant to obesity prevention may not be readily identified through health research approaches. Given a dearth of information to address an urgent public health problem, the information that is available will be valued differently and efforts to fill evidence gaps given high priority.

Study Scope

This study was designed to go beyond obesity prevention as such, also yielding guidance for more general efforts to address complex, multifactorial public health challenges and examining the application of a systems perspective to the contexts in which policy and programmatic decisions that address such challenges are made. Equally important, the study scope was defined to reflect the need for a framework to guide decision making about obesity prevention in the population at large, rather than only in relation to children and youth. The epidemic of obesity in children and youth continues to be a major driver of preventive initiatives, but obesity prevention is also relevant to adults. As noted earlier, most obesity and obesity-related diseases develop during adulthood, and an obese adult population contributes to the development of obesity in children; indeed, from a population perspective, the environments of children and adults are interdependent and inseparable.

The committee defined "framework" to mean a process for identifying, evaluating, and compiling evidence relevant to various obesity prevention decision-making scenarios based on clearly specified concepts, assumptions, or rules. The process of developing the framework involved:

- clarifying what is meant by population-based obesity prevention, that is, the types of initiatives that are undertaken to improve environments for eating and physical activity in ways that will reduce the likelihood of excess weight gain among people in general, as well as among particularly high-risk populations;
- envisioning potential users and uses of such a framework, including the types of decisions to be made and ways in which evidence might enter into these decisions;
- assessing the strengths and limitations of current approaches to evaluating and generating evidence about obesity prevention; and
- identifying research approaches and forms of evidence that could be used to expand the options available to inform decision making on obesity prevention.

The assessment of the strengths and limitations of current approaches included a review of several ongoing efforts to improve research and practice in the field of obesity prevention or public health. These efforts focus on improving methods for characterizing the strength of evidence for prevention approaches, incorporating issues of generalizability of the research or relevance to practice into effectiveness studies, identifying appropriate outcome measures for obesity prevention, or enhancing the reporting of the results of prevention studies. Appendix B describes these efforts.

This report has two primary audiences: (1) decision makers and the intermediaries who assist them in making decisions and (2) because much of the evidence that is needed does not now exist, those who conduct research relevant to obesity prevention or who evaluate existing programs, as well as research-trained academics or other professionals with direct access to and knowledge of methods of evidence review and synthesis. Other important audiences include research funders and publishers of research findings.

OBESITY AS A SOCIETYWIDE PROBLEM

Overweight and obesity—reflecting a gain of excess body fat—result from the cumulative and initially imperceptible effects of hour-to-hour and day-to-day eating and physical activity behaviors that create a surplus of calories consumed relative to calories expended. Excess weight gain is usually unintentional and gradual, and it is difficult to reverse on a permanent basis.

Obesity prevention includes the need both to help those who are already overweight or obese avoid further weight gain and to help people who are not overweight or obese to avoid gaining excess weight. As suggested above, even if all those who are currently obese were to be treated effectively today, growth in the numbers of people becoming obese would continue without preventive efforts. Trend data from the National Health Examination and the National Health and Nutrition Examination Surveys between 1960 and 2002 indicate that the average body weight of U.S. men and women increased by 24.7 pounds (11.2 kg) and 24.1 pounds (11.0 kg), respectively, while there was a relatively small increase in average height. The associated shift in average body mass index (BMI)—the typical weight-for-height measure used to assess obesity—was 3 kg/m^2 in adults, in both men and women, from approximately 25 to 28 (Ogden et al., 2004). Figure 1-1 shows the shift in the distribution of BMI among U.S. adults between 1976-1980 and 2005-2006, with the gain being most pronounced among the heaviest people (Ogden et al., 2007).

Data from the 2007-2008 National Health and Nutrition Examination Surveys (NHANES)[2] show that among adults aged 20 or older, nearly 34 percent have weight levels in the obese range (BMI ≥ 30) and another 34 percent are classified as overweight (BMI between 25 and 30), thus the combined prevalence of overweight and obesity is nearly 68 percent (Flegal et al., 2010). Among children and adolescents aged 2 through 19, nearly 17 percent are classified as obese (BMI for age and sex ≥ 95th percentile), and 15 percent are classified as overweight (BMI for age and sex ≥ 85th but < 95th percentile), thus close to 32 percent are either overweight or obese (Ogden et al., 2010). There is no evidence that underlying biological susceptibility to weight gain has changed, although there is ample evidence of increases in such factors as the amount of food available and where it is available, creating eating environments that are highly conducive to the consumption, often unintentional, of excess calories (Kral and Rolls, 2004; Ledikwe et al., 2005; Story et al., 2008; Wansink, 2004). There is also no evidence that the prevalence of individual weight control attempts has decreased as average population weights have increased. For example, the proportion of U.S. adult men and women who reported attempting to lose weight was, respectively, 29 and 44 percent in 1996 and 33 and 46 percent in 2000 (Bish et al., 2005; Serdula et al., 1999). Accordingly, the increasing prevalence of obesity in diverse populations reveals the limitations of biological and behavioral controls on caloric balance in the face of high caloric availability and the relatively low levels of physical activity that are common in urban and westernized societies (Booth et al., 2001; French et al., 2001; IOM, 2005; Kumanyika et al., 2002; Popkin and Gordon-Larsen, 2004). The need for multifaceted actions to prevent obesity should no longer be a subject of serious debate.

Many multifaceted interventions for obesity prevention influence the calorie balance by focusing on either energy intake or energy expenditure. According to Swinburn and colleagues (2009), it is important to know whether one of these two factors is the primary driver of the obesity epidemic in order to prioritize potential interventions. The relative contributions of increased energy intake and decreased energy expenditure to the obesity epidemic have been debated, with some suggesting that increased energy intake is sufficient to explain the weight gain of the American population since the 1970s (Bleich et al., 2008; Swinburn et al., 2009). Both factors are clearly important for energy balance and other health outcomes, and this report does not aim to tell decision makers whether to focus on energy intake, energy expen-

[2] NHANES is a continuous program of studies designed to assess the health and nutritional status of a nationally representative sample of children and adults in the United States.

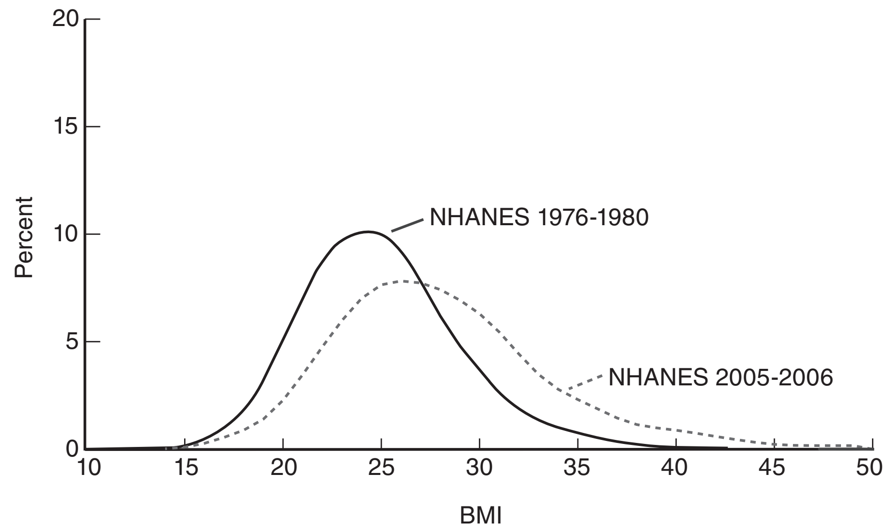


FIGURE 1-1 Changes in the distribution of body mass index (BMI) between 1976-1980 and 2005-2006 among U.S. adults aged 20-74.
NOTE: NHANES = National Health and Nutrition Examination Survey, a continuous program of studies designed to assess the health and nutritional status of a nationally representative sample of children and adults in the United States.
SOURCE: Ogden et al., 2007.

diture, or both. Indeed, the relative importance of each may vary across populations and settings.

Figures 1-2 and 1-3 show recent levels of obesity in three racial/ethnic groups in the United States and projections of future prevalence under the assumption that recently observed rates of increase will continue (Wang et al., 2008). The approximately one-third of men in all three groups who were obese in 1999-2004 could increase to nearly one-half within two decades. The same could be true for non-Hispanic white women, with even higher levels for non-Hispanic black and Mexican American women, for whom the prevalence of obesity was already greater than 50 percent and 40 percent, respectively.

Prevention efforts should not only aim to improve weight levels in the population at large but also target subpopulations with higher-than-average risk. Figures 1-2 and 1-3 identify key high-risk population subgroups—Mexican Americans and African Americans (Hedley et al., 2004; Ogden et al., 2007). Other U.S. ethnic groups with above-average obesity prevalence include American Indians, Pacific Islanders, and Native Hawaiians (Anderson and Whitaker, 2009; Broussard et al., 1991; Davis et al., 2004; Kumanyika, 1993). Low-income children and youth are also at high risk for obesity.

In many cases, above-average obesity prevalence is also observed in low-income populations, or high income is associated with lower obesity prevalence. In adults, the inverse gradient in obesity levels is seen most clearly in white females (Ogden, 2009). Figure 1-4 shows data for children with both sexes combined (Freedman et al., 2007).

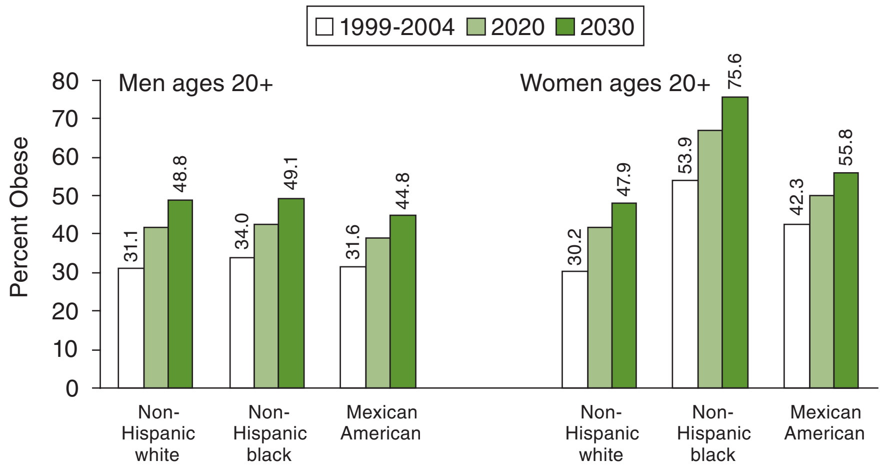


FIGURE 1-2 Projected increases in adult obesity (defined as BMI ≥ 30 kg/m^2) if current trends continue, for males and females in three U.S. ethnic groups.
SOURCE: Adapted from Wang et al., 2008.

Obesity prevalence decreases with increasing income in non-Hispanic white children and is lowest in the high-income group among Mexican American children. No clear income gradient is observed in African American children, although higher risk with increasing income is suggested. However, the proportion of African American children who are obese is higher than that of white children at all income levels. The overall burden of obesity associated with living below the poverty line applies to 43 percent of African American and 40 percent of Mexican American children and to 16 percent of white children.

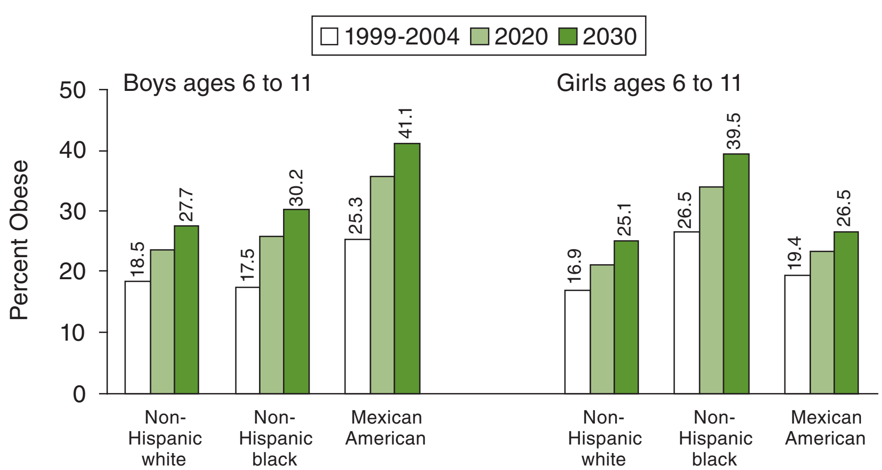


FIGURE 1-3 Projected increases in childhood obesity (defined as BMI ≥ 95th percentile based on CDC growth charts) if current trends continue, for males and females in three U.S. ethnic groups.
SOURCE: Adapted from Wang et al., 2008.

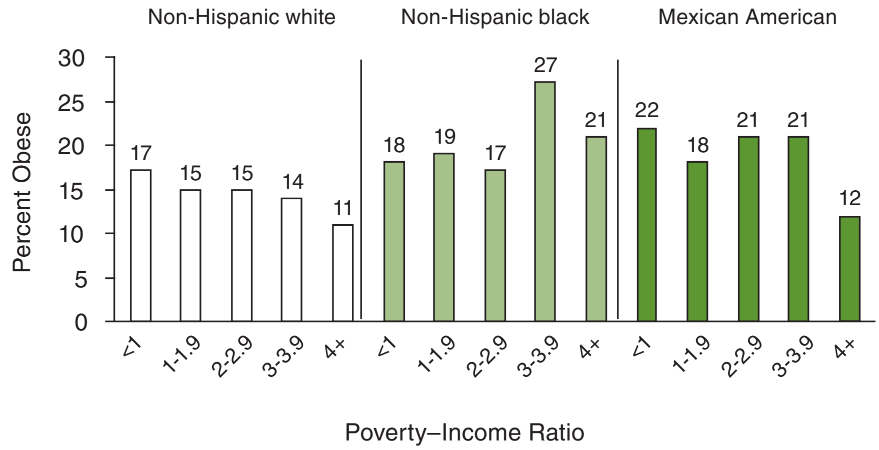


FIGURE 1-4 Percent of children and youth aged 2-19.9 who are obese (defined as BMI ≥ 95th percentile based on CDC growth charts), by family poverty–income ratio.
NOTES: The poverty–income ratio is the ratio of the income of the family to family income at the poverty level. Families with an income ratio of < 1 are below the poverty threshold.
SOURCE: Freedman et al., 2007.

While obesity is a priority from an epidemiologic and health status perspective, it gains further precedence because of the way it influences other aspects of American society. Obesity carries substantial direct and indirect costs for the nation's economy, such as discrimination, economic disenfranchisement, lost productivity, and disability. As a result, states and communities end up diverting resources to prevention and treatment, and the nation's health care system is burdened with the comorbidities of obesity, such as type 2 diabetes, hypertension, cardiovascular disease, osteoarthritis, and cancer (IOM, 2005). According to a recent analysis (Finkelstein et al., 2009), the annual medical burden of obesity is nearly 10 percent of all medical spending and may have risen to $147 billion per year by 2008. Further, obesity places at risk the long-term welfare and readiness of the U.S. military services by reducing the pool of individuals eligible for recruitment and decreasing the retention of new recruits (IOM, 2005). In addition to these societal costs, obesity imposes a broad range of costs on individuals. Tables 1-1 and 1-2, respectively, highlight the myriad effects of adult and child obesity on physical and mental health and quality of life.

DECISION MAKING, OBESITY PREVENTION, AND EVIDENCE NEEDS

In an ideal world, perhaps, policy decisions would be made primarily as a result of review and analysis of a carefully selected body of evidence that provided certainty and specificity as to the best approach or approaches, the outcomes that would result, and the cost of achieving those outcomes. However, real-life experience demonstrates that policy or programmatic decision making is not always such a linear process and that initiatives grounded in an idea generated by a research study may be the exception. Formal evidence may be used in highly variable ways, at different stages of a

TABLE 1-1 Relative Risk of Health Problems Associated with Obesity in Adults[a]

Greatly Increased (relative risk > 3)	Moderately Increased (relative risk 2-3)	Slightly Increased (relative risk 1-2)
• Type 2 diabetes	• Coronary heart disease	• Cancer[b]
• Gallbladder disease	• Hypertension	• Reproductive hormone abnormalities
• Dyslipidemia	• Osteoarthritis	• Polycystic ovary syndrome
• Insulin resistance	• Hyperuricemia and gout	• Lower back pain
• Breathlessness		• Anesthesia complications
• Sleep apnea		• Fetal defects

[a] All relative risk values are approximate.
[b] Breast cancer in postmenopausal women, endometrial cancer, colon cancer.
SOURCE: Reprinted with permission from World Health Organization, 2000. Copyright (2000) by the World Health Organization.

process of instituting policies or programs, and for different reasons. Decision makers (i.e., those who make a decision to implement an intervention) may consider research-based evidence useful for supporting or opposing certain actions, but may also rely on a range of other information related to feasibility, political acceptability, and public support. Community programs may view research-based evidence as interesting but not entirely relevant to initiatives they may undertake in the real world. Program evaluators may share this view, emphasizing research that begins with projects of proven practicability that are shown to have promise or impact based on how things transpire. Members of the public may be frustrated with new evidence that appears daily and is often discordant with evidence from the day before. Additionally, making decisions about prevention is inherently complex, especially for problems, such as obesity, with multiple types and layers of causation. Moreover, prevention (keeping something

TABLE 1-2 Physical, Social, and Emotional Health Consequences of Obesity in Adults, Children, and Youth

Physical Health	Social Health	Emotional Health
• Cardiovascular disease	• Stigma	• Low self-esteem
• Cancer	• Negative stereotyping	• Negative body image
• Glucose intolerance and insulin resistance	• Discrimination	• Depression
• Type 2 diabetes	• Teasing and bullying	
• Hypertension	• Social marginalization	
• Dyslipidemia		
• Hepatic steatosis		
• Cholelithiasis		
• Sleep apnea		
• Menstrual abnormalities		
• Impaired balance		
• Orthopedic problems		

SOURCES: IOM, 2005; WHO, 2000.

from happening) may be a much less dramatic or salient outcome than treatment (removing a problem), whose effects are not only obvious but also potentially more rewarding—personally and politically. Thus, although the case for prevention is compelling, making the decision to give it priority and determining which strategies will lead to more "nonevents" are highly challenging.

The widely held conviction that such decisions should be "evidence-based" only increases the challenge. How does one best determine which aspects of the environment to change, what changes to make, and how to effect these changes? Which interventions can potentially be effective on a large scale and sustainable over time? Where are the synergies among interventions? Could there be undesirable effects not readily apparent at the outset? Ideally, there would be a firm and comprehensive evidence base to inform the myriad decision makers whose actions—intentionally or unintentionally—influence the social and environmental determinants of unwanted weight gain and obesity. In reality, the approaches to evidence that apply to decision making about the treatment of obesity or other clinical problems are inadequate and sometimes inappropriate for application to decisions about public health initiatives such as obesity prevention. This situation poses a dilemma, one that is clearly compounded by the urgency of deciding which interventions can best address the problem. Evidence is needed to support such decisions for many reasons:

- to justify interventions generally, particularly when there is competition for resources or opposition;
- to inform priority setting and justify actions targeting high-risk populations, differentiate actions that are likely to be effective from those that are not, and quantify likely impacts;
- to estimate costs and cost-effectiveness and anticipate unintended consequences, including harms to individuals or businesses; and
- to help understand implementation factors, that is, what to do and how to do it.

The purpose of this report is to present a framework that will make it possible to meet these evidence needs.

Figure 1-5, which illustrates the levels and sectors of influence on obesity in populations, is helpful for understanding why decisions about obesity prevention are so complex (IOM, 2007). The figure shows the variety of levels at which interventions may be aimed and demonstrates how, taken together, a set of actions across levels might interrelate. Actions within many of these levels are focal points for discussion on how to combat the obesity epidemic within the United States and globally. Pathways or logic models can be drawn to link these levels to weight-related behaviors of population groups and individuals, such as food purchases; intakes of sugar-sweetened beverages, fruits and vegetables, whole grains, and total calories; television watching and other screen time; automobile use; breastfeeding; and routine and recreational physical activity. Relevant initiatives might focus on international trade

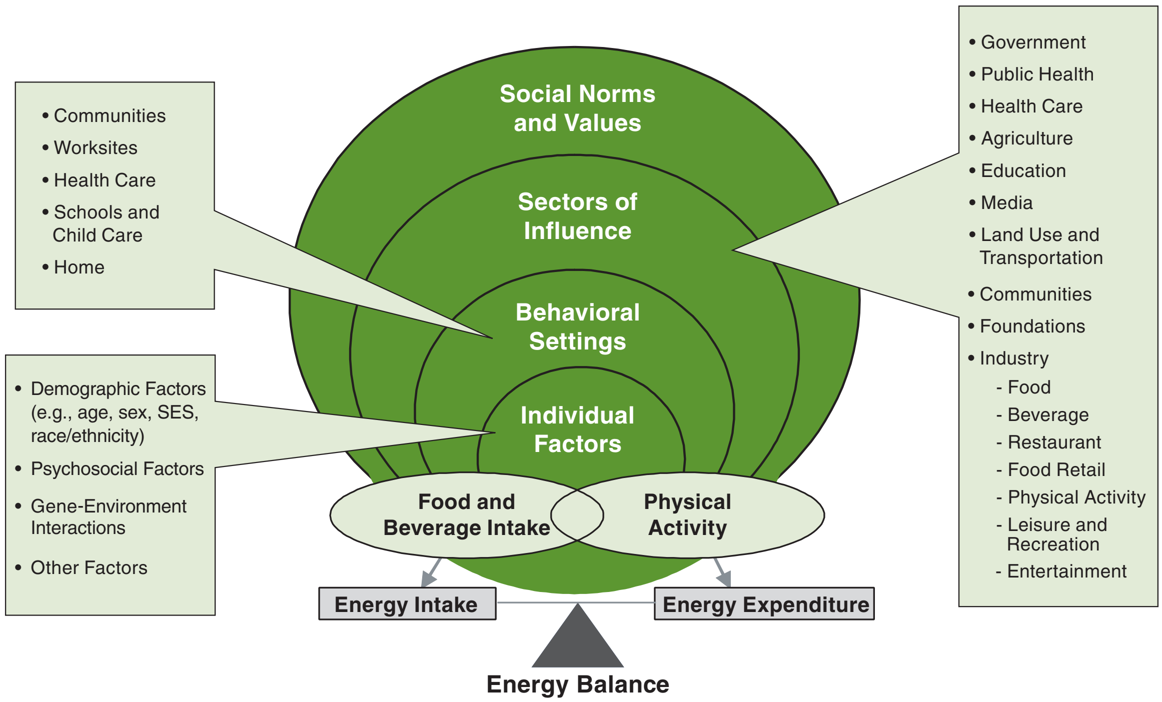


FIGURE 1-5 Comprehensive approach for preventing and addressing childhood obesity.
NOTE: SES = socioeconomic status.
SOURCE: IOM, 2007.

agreements that affect the food supply, food advertising, retail and restaurant food locations and marketing practices (including labeling and portion sizes), food prices, agricultural policies and food production systems, community design aspects that influence outdoor activity, building design aspects that influence indoor activity, public transportation, insurance coverage for wellness services, poverty levels, funding levels for and regulations guiding federal nutrition assistance programs, and school and worksite environments related to food and physical activity (IOM, 2009; James et al., 2006; WHO, 2004; World Cancer Research Fund and American Institute for Cancer Research, 2009).

Obesity prevention initiatives related to the settings and pathways shown on Figure 1-5 are ongoing, and more are undertaken or proposed every day. Many are quite far-reaching in their potential implications for such variables as food prices; food availability; community infrastructure; entertainment; and the nature and flow of information about food, activity, and weight. When the needed evidence exists to inform these decisions, the challenge is how to ensure that it appears in the right place at the right moment in the decision-making process and is taken to heart. If the evidence that exists is faulty or biased, the challenge is how to prevent this evidence from being used to support decisions that will ultimately prove to be poor ones. If the evidence is inadequate, one challenge is to prevent the lack of evidence from precluding timely and needed best guesses about potentially helpful solutions; another is to create the evidence to fill the gap. In the dynamic context of such issues as the obesity epidemic, the uneasy but necessary balance among making use of the best available evidence, avoiding the pitfalls of poor evidence, taking reasonable actions in the face of absent or grossly insufficient evidence, and waiting for the long process of formulating and generating relevant new evidence is a key challenge facing decision makers. Guidance is needed to keep a fruitful balance in view as progress toward solutions is made.

Box 1-1 presents some decision scenarios related to obesity prevention policies and programs. Those shown in italics are decisions likely to be made based on priorities unrelated to obesity but with potential effects on eating or physical activity. The decision scenarios in Box 1-1 all relate in some way to decisions about interventions; these types of questions are central to obesity prevention. However, decisions about the selection, implementation, and evaluation of interventions are closely tied to questions about why the intervention is needed and how a specific intervention can be expected to work in a given context. Some of these scenarios illustrate the potential for making arguments on obesity issues to justify changes in well-established policies and practices that affect marketing, services, or infrastructure related to patterns of eating and physical activity so as to ultimately influence consumer behavior. Consideration of these types of scenarios was core to the committee's deliberations, undertaken to ensure that the resulting framework would be practical and broadly relevant.

ORGANIZATION OF THE REPORT

The remainder of this report is organized as follows:

- Section I provides background on obesity prevention concepts and approaches (Chapter 2) and on evidence issues, including the results of the committee's review of the limitations of current approaches to evaluating obesity prevention research (Chapter 3). Chapter 3 also introduces the committee's proposed framework for using evidence to inform decision making on obesity prevention.
- Section II presents an elaboration of each component of the proposed framework (Chapters 4 through 8). The importance of taking a systems perspective and how such a perspective can be useful in the process of using evidence to inform obesity prevention decisions is discussed in Chapter 4. Chapter 5 presents a typology of decisions and evidence that can guide the task of identifying useful evidence, a classification and discussion of various forms of evidence, and guidance on the process of searching for evidence. Chapter 6 explains the committee's recommended approach to evidence evaluation as embodied in the framework; the discussion includes references to resources that can be useful in assessing the credibility of different forms of evidence. Chapter 7 addresses the

Box 1-1
Decision Scenarios Related to Obesity Prevention Policies and Programs

Decisions about specific, populationwide policy changes with a potential direct impact on physical activity or weight-related behavior of adults or children	• A local board of health or city council considers a proposal to require that menus of chain restaurants show the calorie content of each item. • A state legislator decides whether to vote for or against a measure to tax sugar-sweetened beverages as an obesity prevention measure. • *A city planning authority is considering closing city center streets to automobile use to reduce traffic flow and increase pedestrian access to local businesses.* • A federal authority is attempting to determine whether there are sufficient grounds to implement statutory restrictions on television and Internet advertisements directed at children. • A school board member decides whether to vote for a measure to send parents reports of their children's BMI levels based on annual weight and height screening.
Decisions to undertake or retain a specific program that would affect access to food or physical activity options	• A school board is considering a policy that would remove sugar-sweetened beverages and snack foods from schools in its district. • The director of a youth service organization is about to select a family-oriented nutrition and fitness program for nationwide implementation. • *A state education authority is faced with a decision to cancel the requirement for daily physical education classes to make room for additional academic content in the curriculum.* • A local foundation is considering a proposal to fund a program to create a network of farmers' markets in low-income neighborhoods. • *An economically strapped city is considering cuts in funding for local parks and recreation centers.* • A community development corporation has been asked to partner in an initiative to establish a supermarket in a low-income neighborhood. • The state comptroller may have to cut funding for school bus service in certain neighborhoods. • *A prior mayor instituted a program to link after-school programs that focus on academics and sports. The current mayor is trying to decide whether to continue the program.*
Decisions about which intervention(s) to select from a set of potential options to decrease obesity-promoting environmental influences	• A state health commissioner who has been given 5 years to show progress in reducing childhood obesity rates among low-income communities needs to develop a strategic plan. • *An employer is negotiating what package of wellness services will be included in the employee benefit package.* • A community-based organization is seeking ways to increase access to more healthful foods for residents of a rural community. • A state-based obesity prevention coalition is trying to identify advocacy priorities to address childhood obesity. • A budget director is attempting to weigh the potential cost-effectiveness of a taxation strategy and complementary social marketing campaign to reduce population obesity levels.

NOTE: Scenarios shown in italics are decisions likely to be made based on priorities unrelated to obesity but with potential effects on eating or physical activity.

processes of assembling and reporting results, translating evidence to and from practice, deciding on the implications of the information for decision making, and integrating the overall knowledge gained from the evidence review process. Chapter 8 addresses the need for researchers to use the framework to motivate their work and think systematically about how to make progress in the inevitable, ongoing context of inadequate and imperfect evidence, and highlights other approaches to evidence generation for increased use in research on obesity prevention.

- Finally, Section III turns to next steps. Chapter 9 provides guidance for future efforts of communicating, disseminating, refining, and updating the framework as new evidence and new methods emerge. Chapter 10 summarizes the committee's conclusions and presents recommendations for supporting the identification, use, and evaluation of evidence to inform obesity prevention decisions and generate evidence when it is lacking.

For reference throughout the report, a list of acronyms and glossary of terms is provided in Appendix A; a table listing current efforts to improve research and practice in the field of obesity prevention or public health is given in Appendix B; Appendix C contains a table listing current reviews of obesity prevention interventions and strategies; Appendix D provides a list of selected information sources for locating evidence; Appendix E presents an in-depth discussion of some study design alternatives to randomized controlled trials and how theory, professional experience, and local wisdom can be used when evidence is lacking; Appendix F contains the agenda of two public workshops held by the committee; and Appendix G presents biographical sketches of the committee members.

REFERENCES

Anderson, S. E., and R. C. Whitaker. 2009. Prevalence of obesity among US preschool children in different racial and ethnic groups. *Archives of Pediatrics and Adolescent Medicine* 163(4):344-348.

Bish, C. L., H. M. Blanck, M. K. Serdula, M. Marcus, H. W. Kohl III, and L. K. Khan. 2005. Diet and physical activity behaviors among Americans trying to lose weight: 2000 behavioral risk factor surveillance system. *Obesity Research* 13(3):596-607.

Bleich, S., D. Cutler, C. Murray, and A. Adams. 2008. Why is the developed world obese? *Annual Review of Public Health* 29:273-295.

Booth, S. L., J. F. Sallis, C. Ritenbaugh, J. O. Hill, L. L. Birch, L. D. Frank, K. Glanz, D. A. Himmelgreen, M. Mudd, B. M. Popkin, K. A. Rickard, S. St. Jeor, and N. P. Hays. 2001. Environmental and societal factors affect food choice and physical activity: Rationale, influences, and leverage points. *Nutrition Reviews* 59(3, Supplement 2):S21-S39.

Broussard, B. A., A. Johnson, J. H. Himes, M. Story, R. Fichtner, F. Hauck, K. Bachman-Carter, J. Hayes, K. Frohlich, N. Gray, S. Valway, and D. Gohdes. 1991. Prevalence of obesity in American Indians and Alaska natives. *American Journal of Clinical Nutrition* 53(6, Supplement).

Davis, J., J. Busch, Z. Hammatt, R. Novotny, R. Harrigan, A. Grandinetti, and D. Easa. 2004. The relationship between ethnicity and obesity in Asian and Pacific Islander populations: A literature review. *Ethnicity and Disease* 14(1):111-118.

Finkelstein, E. A., J. G. Trogdon, J. W. Cohen, and W. Dietz. 2009. Annual medical spending attributable to obesity: Payer and service-specific estimates. *Health Affairs* 28(5): w822-w831.

Flegal, K. M., M. D. Carroll, C. L. Ogden, and L. R. Curtin. 2010. Prevalence and trends in obesity among U.S. adults, 1999-2008. *Journal of the American Medical Association* 303(3):235-241.

Freedman, D. S., C. L. Ogden, K. M. Flegal, L. K. Khan, M. K. Serdula, and W. H. Dietz. 2007. Childhood overweight and family income. *Medscape General Medicine* 9(2).

French, S. A., M. Story, and R. W. Jeffery. 2001. Environmental influences on eating and physical activity. *Annual Review of Public Health* 22:309-335.

Hedley, A. A., C. L. Ogden, C. L. Johnson, M. D. Carroll, L. R. Curtin, and K. M. Flegal. 2004. Prevalence of overweight and obesity among U.S. children, adolescents, and adults, 1999-2002. *Journal of the American Medical Association* 291(23):2847-2850.

IOM (Institute of Medicine). 2005. *Preventing childhood obesity: Health in the balance*. Edited by J. Koplan, C. T. Liverman, and V. I. Kraak. Washington, DC: The National Academies Press.

IOM. 2006. *Food marketing to children and youth: Threat or opportunity?* Edited by J. M. McGinnis, J. Gootman, and V. I. Kraak. Washington, DC: The National Academies Press.

IOM. 2007. *Progress in preventing childhood obesity: How do we measure up?* Edited by J. Koplan, C. T. Liverman, V. I. Kraak, and S. L. Wisham. Washington, DC: The National Academies Press.

IOM. 2009. *Local government actions to prevent childhood obesity.* Edited by L. Parker, A. C. Burns, and E. Sanchez. Washington, DC: The National Academies Press.

James, W. P. T., N. J. Rigby, R. J. Leach, S. Kumanyika, T. Lobstein, and B. Swinburn. 2006. *Global strategies to prevent childhood obesity: Forging a societal plan that works*. Paper presented at Global Prevention Alliance McGill Integrative Health Challenge, October 26-27, 2006.

Kral, T. V. E., and B. J. Rolls. 2004. Energy density and portion size: Their independent and combined effects on energy intake. *Physiology and Behavior* 82(1):131-138.

Kumanyika, S. K. 1993. Special issues regarding obesity in minority populations. *Annals of Internal Medicine* 119(7, Part 2):650-654.

Kumanyika, S., R. W. Jeffery, A. Morabia, C. Ritenbaugh, and V. J. Antipatis. 2002. Obesity prevention: The case for action. *International Journal of Obesity* 26(3):425-436.

Ledikwe, J. H., J. A. Ello-Martin, and B. J. Rolls. 2005. Portion sizes and the obesity epidemic. *Journal of Nutrition* 135(4):905-909.

Ogden, C. L. 2009. Disparities in obesity prevalence in the United States: Black women at risk. *American Journal of Clinical Nutrition* 89(4):1001-1002.

Ogden, C. L., C. D. Fryar, M. D. Carroll, and K. M. Flegal. 2004. *Mean body weight, height, and body mass index, United States 1960-2002*. Advance data from Vital and Health Statistics, No. 347. Hyattsville, MD: National Center for Health Statistics.

Ogden, C. L., M. D. Carroll, M. A. McDowell, and K. M. Flegal. 2007. *Obesity among adults in the United States—No statistically significant change since 2003-2004*. NCHS Data Brief No. 1. Hyattsville, MD: National Center for Health Statistics.

Ogden, C. L., M. D. Carroll, L. R. Curtin, M. M. Lamb, and K. M. Flegal. 2010. Prevalence of high body mass index in U.S. children and adolescents, 2007-2008. *Journal of the American Medical Association* 303(3):242-249.

Popkin, B. M., and P. Gordon-Larsen. 2004. The nutrition transition: Worldwide obesity dynamics and their determinants. *International Journal of Obesity* 28(Supplement 3):S2-S9.

Serdula, M. K., A. H. Mokdad, D. F. Williamson, D. A. Galuska, J. M. Mendlein, and G. W. Heath. 1999. Prevalence of attempting weight loss and strategies for controlling weight. *Journal of the American Medical Association* 282(14):1353-1358.

Story, M., K. M. Kaphingst, R. Robinson-O'Brien, and K. Glanz. 2008. Creating healthy food and eating environments: Policy and environmental approaches. *Annual Review of Public Health* 29:253-272.

Swinburn, B., G. Sacks, and E. Ravussin. 2009. Increased food energy supply is more than sufficient to explain the U.S. epidemic of obesity. *American Journal of Clinical Nutrition* 90(6):1453-1456.

Wang, Y., M. A. Beydoun, L. Liang, B. Caballero, and S. K. Kumanyika. 2008. Will all Americans become overweight or obese? Estimating the progression and cost of the U.S. obesity epidemic. *Obesity* 16(10):2323-2330.

Wansink, B. 2004. Environmental factors that increase the food intake and consumption volume of unknowing consumers. *Annual Review of Nutrition* 24:455-479.

WHO (World Health Organization). 2000. *Obesity: Preventing and managing the global epidemic. Report of a WHO consultation*. WHO Technical Report Series 864. Geneva: World Health Organization. Available online: http://www.who.int/nutrition/publications/obesity/en/index.html.

WHO. 2004. *Global strategy on diet, physical activity, and health*. Geneva: World Health Organization.

World Cancer Research Fund and American Institute for Cancer Research. 2009. *Policy and action for cancer prevention. Food, nutrition, and physical activity: A global perspective*. Washington, DC: American Institute for Cancer Research.

2

Obesity Prevention Strategies in Concept and Practice

KEY MESSAGES

- From a population perspective, obesity prevention means lowering the mean body mass index (BMI) level and decreasing the rate at which people enter the upper end of the BMI distribution. Accomplishing this will require that adults at a healthy weight maintain that status and that children maintain a healthy weight trajectory and avoid excess weight gain.
- Obesity prevention takes place in diverse settings that provide access to whole populations or high-risk subpopulations.
- The most promising approaches for obesity prevention are population-based and multilevel, focus on environmental and policy change, and require participation from actors in multiple sectors.
- The strategies and actions undertaken to prevent obesity and their resulting outcomes vary according to the different environments in which they are undertaken.

Obesity has long been viewed as an individual problem, treated primarily by individual or family-based medical and/or psychological methods. As noted in Chapter 1, the expansion of the obesity epidemic to affect diverse age groups, socio-economic strata, and racial or ethnic groups and communities in the United States and globally has caused a paradigm shift toward preventive strategies that can reach whole communities and populations, as well as a sense of urgency in mounting these strategies (Kumanyika et al., 2002, 2008; Public Health Service, 2001; WHO, 2000). This chapter provides an overview of current thinking and practice regarding obesity prevention strategies. First, definitions and types of obesity prevention are discussed from a population perspective. The logic of obesity prevention when undertaken by different sectors and at different levels is then examined using an overarching model developed by a prior Institute of Medicine (IOM) committee on child obesity prevention. Specific obesity prevention approaches are next identified across a spectrum that includes changes in policy and legislation, organizational practices, community and neighborhood environments, and health care settings, as well as strategies based on health communication or health education and individual counseling. Examples of how these strategies are combined in multilevel approaches are also provided. The chapter concludes by placing strategies being used to address obesity prevention in the context of those used to address other complex public health problems.

Approaches discussed in this chapter draw on accounts of current practice, policies, and programs that have been implemented or recommended as being potentially pertinent to obesity prevention. Background information on concepts and evidence for assessing the effectiveness of the approaches presented here is provided in Chapter 3.

DEFINITIONS AND TYPES OF PREVENTION

The concept of increasing obesity prevalence caused by a shift in the population distribution of BMI was shown in Figure 1-1 in Chapter 1, based on actual data for adults in the U.S. population. From a population perspective, obesity prevention means reversing this shift—lowering the mean BMI level and decreasing the incidence of obesity, that is, the rate at which people enter the upper end of the BMI distribution (Kumanyika et al., 2002; WHO, 2000). Stabilization of the BMI distribution or the incidence of new people crossing the relevant BMI threshold (30 kg/m^2 for adults or the 95th sex–age percentile of the BMI reference for children) would be the first sign of success, followed by a decline in these indicators. Within the population, accomplishing this will require that adults at a healthy weight (i.e., not underweight, overweight, or obese) maintain that status and avoid the phenomenon of excess weight gain that commonly accompanies aging. For children, prevention means maintaining a healthy weight trajectory and preventing excess weight gain while growing, developing, and maturing. Prevention for individuals who are already overweight or obese can be defined as avoiding progression to a more severe level of obesity and/or preventing or delaying the onset of obesity-related medical conditions (IOM, 1995).

A 1995 IOM report on criteria for evaluating weight management programs emphasizes the importance of giving greater priority to obesity prevention in light of the increases in population weight levels and obesity prevalence (IOM, 1995). Although that report focuses primarily on setting standards for individually focused weight management programs, the committee that produced it set the stage for thinking about levels of prevention from a population perspective. Instead of the *primary*, *secondary*, and *tertiary* prevention terminology, based on a continuum of disease emergence,[1] that is typically used in relation to chronic diseases, the committee identified categories of prevention in terms of population segments. Thus, the 1995 committee moved toward a more general risk reduction approach to prevention.

As shown in Figure 2-1, the population-based classification in the 1995 report identifies three levels of obesity prevention interventions targeting the population at large (universal); subgroups of the population that are at high risk, such as those defined by race/ethnicity or income level (selective); or individuals at high risk (indicated). High risk may refer to individuals who are identified as having minimal but detectable signs that foreshadow obesity, who exhibit biological markers indicating a predisposition to an obesity-related condition, or who do not meet the full diagnostic criteria for the disorder itself (IOM, 1995). The key distinction is in whether the population targeted has been selected on the basis of risk characteristics. The adoption of a similar population-based conceptualization of obesity prevention by a 1997 World Health Organization (WHO) expert panel in characterizing intervention strategies to address the obesity epidemic globally (WHO, 2000) has further disseminated this way of thinking about obesity prevention strategies.

This population-based view of prevention focuses attention on key issues related to how resources are used and the relative emphasis placed on interventions aimed at directly changing individual behavior (e.g., in indicated prevention). Targeted, selective, and indicated prevention approaches differ in cost per person reached and in many other variables that determine their feasibility and ultimate impact. Fundamental to the need for the new framework proposed in this report, these strategies also differ in how one goes about evaluating their impact—by examining effect sizes, certainty of effects, and potential benefits and harms. Furthermore, a mix of strategies is needed. Individual-oriented strategies are critical, but for problems affecting whole populations, they should be used judiciously because they are often some of the most resource-intensive approaches. Such indicated prevention strategies are essentially treatment strategies directed either at adults or children who can be considered overweight ("pre-obese") or at obese children to prevent their progression to obesity during adulthood. These strategies require case identification and intensive, long-term counseling.

Another reason to place less emphasis on individual-oriented strategies for obesity prevention relates to the nature of what constitutes prevention from an individual behavioral perspective. Although the prevention and treatment of obesity both rely on the same principle of energy balance, the challenges are quite different. Treatment relies on the creation of negative energy balance through substantial reductions in caloric intake and increases in physical activity to produce an observable reduction in body weight. Prevention strives to create caloric balance (or caloric intake that supports appropriate growth patterns in children) and to avoid a persistent, small positive

[1] Interventions implemented before clinically significant disease is evident are termed *primary prevention* (decreasing the development of new cases); *secondary prevention* interventions focus on slowing the progression of the disease and preventing the development of complications; and *tertiary prevention* mitigates disease consequences.

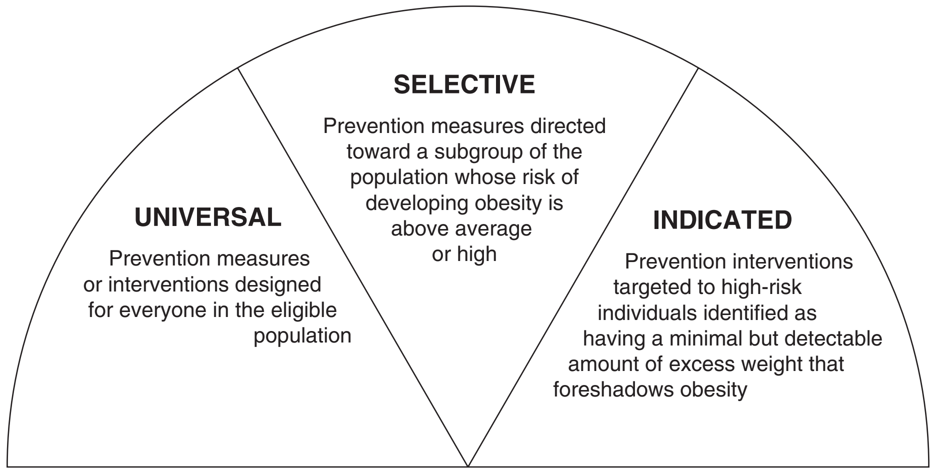


FIGURE 2-1 Types of obesity prevention interventions.
SOURCE: IOM, 1995.

energy imbalance that over time results in excess weight gain. As a result, quantifying the results of prevention is difficult because the results are less apparent—*no change* in weight or health outcomes, as opposed to the weight loss and decrease in risk factors associated with treatment (Kumanyika et al., 2008). It is also difficult to measure or, from the individual's perspective, to be reassured that people are making the small changes in caloric intake and physical activity that may facilitate the prevention of excess weight gain. For this reason, it is important to emphasize universal prevention strategies intended to foster aggregate behaviors along lines that will decrease the likelihood of chronic excess caloric consumption in relation to expenditure. Such approaches focus primarily on modifying social norms and social structures through changes in policies or the implementation of broad interventions that reduce exposure to obesity-promoting environments in the population at large. The same is true for targeted prevention approaches that direct universal strategies at subpopulations (e.g., children and youth, or ethnic or socioeconomic status groups with higher-than-average obesity prevalence). A focus on specific populations may be needed to ensure that interventions address disparities in social and environmental conditions related to eating or physical activity (e.g., in food availability or affordability or in options for physical activity). Universal prevention may address individuals, although not on a one-to-one basis. Examples are mass education or social marketing strategies to improve individual knowledge and skills related to weight change or to promote favorable responses to changes at the environmental or policy level.

Complementary to concepts of obesity prevention that are based on the population segment targeted, public health strategies may also be characterized as affecting macro versus micro environments: macro environments might be at the national, state, or community level, involving variables that affect individuals indirectly, while micro environments might be school, worksite, clinical, or home environments to which individuals are directly exposed on a day-to-day basis. A distinction also is often made between "upstream" and "downstream" approaches: the former address factors far removed from the individual's control, whereas the latter reach individuals on a one-to-one basis. Upstream approaches are also more likely than downstream approaches to reduce disparities for socially and economically disadvantaged populations because they improve access to opportunities for healthful eating and physical activity (Kumanyika et al., 2008). Upstream approaches often target sectors outside of public health and health care, such as education, agriculture, transportation, urban planning, and industry, that influence eating and activity patterns.

THE LOGIC OF OBESITY PREVENTION

To develop a framework for evaluating childhood obesity efforts, a prior IOM committee developed a series of logic models that are helpful in understanding pathways to obesity prevention at different levels and in different sectors. The overall model developed by that committee, which is relevant to obesity prevention for populations of all ages, is reproduced as Figure 2-2 to provide a conceptual basis for thinking about prevention strategies.

At the far right of this logic model, the ultimate outcomes related to obesity prevention ("Health Outcomes") are shown to result from achievements related to four different types of intermediate outcomes—structural, institutional, systemic, and envi-

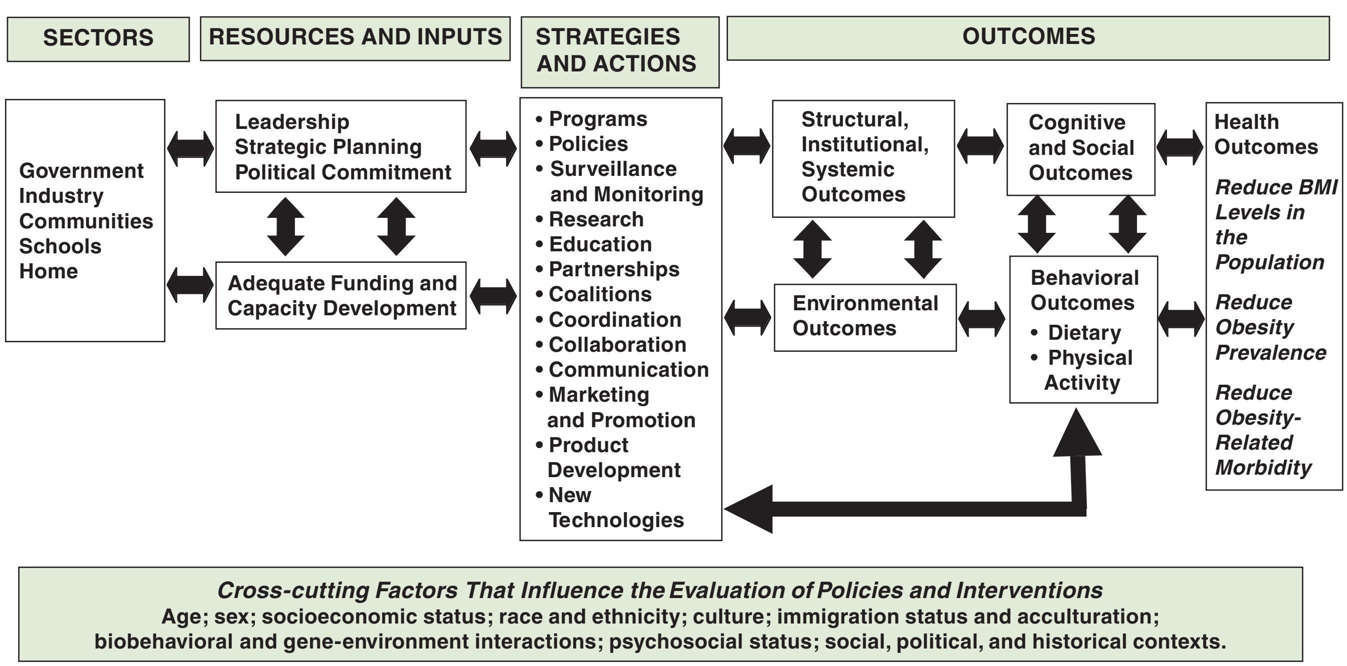


FIGURE 2-2 Evaluation framework for childhood obesity prevention policies and interventions.
SOURCE: IOM, 2007.

ronmental outcomes in society or the community—which impact cognitive, social, and behavioral outcomes in individuals. A range of potential strategies and actions can be applied to achieve the intermediate outcomes.

At the left of the model, the potential for these strategies and actions to be undertaken by various sectors (government, industry, communities, schools, and homes) is indicated. The specifics of resources and inputs, strategies and actions, and outcomes that are relevant vary by sector. How this logic plays out sector by sector is included in the prior committee's report (IOM, 2007). The demographic, sociocultural, and economic variables shown across the bottom of the figure as "cross-cutting factors" might define considerations for when targeted prevention is needed. The individual cross-cutting factors will determine the heterogeneity of effects on individuals and, therefore, the average effects on populations.

The logic model in Figure 2-2 illustrates some key points that will be revisited throughout the remaining chapters of this report. One is that the ultimate objective of obesity prevention activities is a change in BMI levels. Another is that change in BMI outcomes is achieved through activities that have an impact on four types of intermediate outcomes. The model emphasizes that many strategies undertaken to prevent obesity will not be expected to have an impact directly on BMI, but rather on pathways that will alter the context in which eating, physical activity, and weight control occur. This point has clear implications for the evaluation of evidence from obesity prevention studies with respect to how well the outcomes assessed match the actions in question. This report does not aim to give specific guidance as to which outcomes to assess; rather, it encourages users of the framework to think broadly when choosing outcomes, considering intermediate outcomes that lead to the ultimate objective of achieving and/or maintaining appropriate BMI levels.

Expert panels concerned both with dietary guidance generally and with the prevention of specific diseases emphasize the importance of obesity prevention (CDC, 2009; HHS, 2008; USDA/HHS, 2005; WHO, 2003, 2004; World Cancer Research Fund and American Institute for Cancer Research, 2007). For example, the 2005 *Dietary Guidelines for Americans* include a recommendation to "maintain body weight in a healthy range" and to "prevent gradual weight gain over time" (USDA/HHS, 2005, p. 14). The World Cancer Research Fund 2007 report (p. 373) includes a recommendation to "be as lean as possible within the normal range [defined as the appropriate ranges issued by national governments or WHO] of body weight." These groups specify targets for obesity prevention based on population behaviors that may predispose to excess caloric intake or inadequate physical activity. These targets, supported by evidence with varying degrees of certainty, are summarized in Box 2-1. With a few exceptions, these targets apply to people of all ages and are therefore appropriate for both universal and targeted prevention strategies.

Box 2-1
Target Behaviors for Obesity Prevention

Diet-related

- Increase the consumption of fruits and vegetables[a,b] and legumes, whole grains, and nuts.[c]
- Limit calories from added sugars, solid fats, and alcohol.[b,d,e]
- Decrease the consumption of sugar-sweetened beverages[a]/soft drinks.[b]
- Increase breastfeeding initiation, duration, and exclusivity.[a,b,d]
- Reduce the consumption of high-calorie, energy-dense foods.[a,b,c,d]
- Mothers accept their child's ability to regulate energy intake rather than eating until the plate is empty.[b]
- Ensure appropriate micronutrient intake to promote optimal linear growth.[b]

Activity-related

- Increase physical activity/promote an active lifestyle.[a,b,c,d,e,f]
- Children get at least 60 minutes of physical activity per day; adults aim, per week, for 150 minutes of moderate-intensity or 75 minutes of vigorous-intensity physical activity.[f]
- Decrease television viewing.[a,b]

Other

- Limit exposure of young children to heavy marketing of energy-dense, micronutrient-poor foods.[b]
- Provide information and skills necessary to make healthy food choices.[b]

SOURCES:
[a] CDC, 2009.
[b] WHO, 2003.
[c] WHO, 2004.
[d] World Cancer Research Fund and American Institute for Cancer Research, 2007.
[e] USDA/HHS, 2005.
[f] HHS, 2008.

THE SPECTRUM OF OBESITY PREVENTION APPROACHES

This section describes a sampling of obesity prevention interventions, adapted along lines suggested by the Spectrum of Prevention developed by Cohen and Swift (Cohen and Swift, 1999; Rattray et al., 2002) for use in public health and health promotion. The Spectrum of Prevention is a framework for developing multifaceted approaches to address complex public health problems. It identifies multiple levels of increasing scope, from the individual level to the policy level, emphasizing that effectiveness in addressing a given issue requires integrated actions across those levels. In this sense, the spectrum is similar to ecological models used to organize objectives and intervention approaches (McLeroy et al., 1988; Stokols et al., 1996). Ecological models describe individual, interpersonal, organizational, environmental, and policy factors.

A framework conceptualizing prevention actions based on an ecological model is described in a prior IOM report about obesity (IOM, 2005). The way the ecological model has evolved to give rise to thinking about obesity on a broader systems level is discussed in Chapter 4.

Policy and Legislative Approaches

Increased recognition of obesity and especially childhood obesity as a public health crisis has spurred legislative, regulatory, and policy proposals at the national, state, county, and city levels. States and local communities have been more successful than national government agencies in passing and implementing such proposals (IOM, 2009a,b). Referring to Figure 2-2, these are examples of strategies and actions undertaken by the government sector. In the case of reducing consumption of sugar-sweetened beverages and high-calorie, energy-dense foods, policy and legislative approaches might include limiting the availability of these products in certain environments (such as schools) or programs (such as federal nutrition assistance programs), imposing taxes or price increases on them, or mandating a reduction in their standard serving sizes. In the case of increasing physical activity, examples of policy and legislative approaches are requirements and funding for bike paths, mixed land use, parks, and safe routes to school; physical activity requirements for schools and child care centers; and taxes on gasoline.

For the most part, such population-based initiatives are not "proven." That is, they are not yet supported by a strong evidence base showing that they will bring about improved dietary habits, increased physical activity, or reduced obesity at the population level; rather, they are based on extrapolations from clinical trials, expert opinion, previous experience, or promising practices (Brownson et al., 2006; IOM, 2009a; Khan et al., 2009). Nonetheless, because obesity is epidemic, there has been demand in some communities to take immediate action.

Approaches Involving Changes in Organizational Policies, Environments, and Practices

To reach the population, strategies should be implemented across many organizational settings. Referring to Figure 2-2, these settings correspond to strategies and actions involving programs, coalitions, coordination, and collaboration. Examples of organizational settings include child care facilities for younger children, schools for older children and adolescents, colleges for young adults, worksites for other adults, and senior centers for older adults. In addition, community organizations, faith-based settings, and primary care offices are important across the age spectrum (IOM, 2009b; Story et al., 2008). Relevant policies or programs could involve regulatory (mandatory) changes that affect an organization or voluntary changes that could be implemented by an organization itself. Examples of organizational initiatives include requiring restaurants to post the calorie content of their foods on menus; increasing the availability of fruits and vegetables in a worksite cafeteria; and limiting the availability of sugar-sweetened beverages and high-calorie, energy-dense foods in a school. The relative cost of foods served or sold in worksite cafeterias, schools, or community settings might be adjusted. Requirements for physical education and recess in schools, child care settings, and aftercare settings might be instituted and enforced at the federal or state level, but also at the organizational level. Schools might implement training programs to enable teachers to offer nutrition or physical education, provide financial support for programs and services related to weight control, or keep school recreational facilities open after hours for community use. The appeal of such setting-based approaches includes the ability to work with a "captive audience"; provide age- and developmentally appropriate activities; and influence social norms within the setting, with possible transfer to behavior outside of the setting.

Approaches Designed to Foster Changes in Communities and Neighborhoods

Many obesity prevention efforts are undertaken at the local or community level; those focusing on childhood obesity have tended to concentrate on local, district, or state school policies that impact nutrition and/or physical activity opportunities. Policy changes are often easier at the local level because they may be able to be accomplished without the same intensity of representation of vested interests, or through relatively agile regulatory bodies rather than through legislation (IOM, 2009a,b). According to testimony from the Assistant Commissioner of the New York City Department of Health at a public workshop convened by the committee (January 8, 2009), in deciding what strategies to pursue, the department targeted actions that the city could influence through regulations under its jurisdiction. A single individual, community organization, or stakeholder is also apt to have greater influence or power at the local level because opposition at the state (or national) level is likely to be more organized and better financed. Furthermore, it may be easier to achieve consensus at the local than at the state level. An individual or a few committed people can often move an agenda item forward quickly and play a greater role in ensuring its implementation because of the smaller, more manageable scale. Strategy planners at the local level are also more likely to be sensitive to context—they know what is feasible, practical, and acceptable in a particular setting. Different types of evidence—for example, more locality-specific—may be needed for smaller-scale, locally controlled policies that will not encompass a wide variety of settings. Local actions may, therefore, be less vulnerable to the greater costs and inertia entailed in larger-scale policies.

In many cases, policy initiatives were first implemented at the local or community level before being implemented more broadly across a state. Mirroring the history of tobacco control efforts, this scaling up of implementation has been undertaken in current policy and environmental efforts to improve nutrition, increase physical activity, and reduce sedentary activities (IOM, 2009a; Khan et al., 2009). After a pioneering locality develops an intervention, tests its feasibility, and assesses public support and opposition, and the idea becomes less novel over time, similar interventions are implemented more readily in other settings. In addition, legal challenges may have been part of the process. For example, New York City was not successful in implementing calorie posting on menus on the first try (Mello, 2009) because of a successful legal challenge. The city promulgated an amended calorie labeling law in January 2008, after which the courts ruled in its favor against the New York State Restaurant Association. After the second ruling, other New York counties copied much of the New York City language in their legislative proposals in hopes of avoiding a legal challenge.

Examples of strategies and actions that might be undertaken by communities and neighborhoods to improve dietary behaviors include advocating for zoning that limits fast food establishments, facilitating the situating of supermarkets in underserved neighborhoods (based on observational data that access to a supermarket is associated with better dietary quality and/or increased fruit and vegetable intake [Larson et al., 2009; Lovasi et al., 2009; Moore et al., 2008; Morland et al., 2002; Zenk et al., 2009]), increasing the availability of farmers' markets, and creating and supporting community gardens. Examples of strategies and actions to increase physical activity include supporting safe routes to school, creating and maintaining a walking school bus, increasing the number of parks and bike paths, and enhancing walking infrastructure (IOM, 2009a).

Two recent reports recommend community-based policy and environmental strategies and associated measurements that local governments and communities can use to plan, implement, and monitor population-level initiatives designed to prevent obesity (IOM, 2009a; Khan et al., 2009). For most of the strategies, the evidence is insufficient to determine effectiveness in changing caloric intake, physical activity, and/or obesity prevalence; for some of the strategies, there is no evidence, only theoretical projections. One of the reports (IOM, 2009a) summarizes the strength of the evidence (from intervention or observational studies to related nonresearch sources) and recommends a number of "promising" policy strategies, apart from school-based interventions, based on their anticipated effect and projected reach. Both reports are notable in that the majority of the strategies recommended are not evidence-based because, for the most part, strong evidence that they will achieve the desired outcomes is not available. Nonetheless, these reports, which recommend a number of the same strategies, are a useful starting point for local governments, communities, and organizations seeking guidance on where to focus their efforts. Although not intended as such, they may also be useful guides for federal agencies and/or legislators in their strategic thinking about policies to control obesity.

Health Communication and Social Marketing Approaches

Strategies and actions to prevent obesity can be enhanced through communication, marketing and promotion, and new technologies, as shown in Figure 2-2. Documentaries and television specials have captured the attention of the public and the broader media in pointing out some of the marketing tactics used by the food industry to increase consumption of certain foods. Increased awareness of the health and societal costs associated with excess caloric intake from sugar-sweetened beverages and high-calorie, energy-dense foods has led to social pressure to decrease efforts to encourage overeating and supersizing of meals at some fast food establishments. In addition, government organizations such as CDC and the Produce for Better Health Foundation have developed a campaign promoting fruits and vegetables (Foerster et al., 1995), while CDC sponsored the VERB™ campaign to increase and maintain physical activity among youth (Wong et al., 2004).

Interventions in Health Care Settings

While there is little conclusive evidence that health care interventions can prevent obesity, screening to ensure early detection of diseases/conditions and anticipatory guidance and health promotion counseling are hallmarks of any preventive health care program. For children aged 2-20, annual screening for obesity using sex-specific BMI-for-age percentiles, accompanied by preventive counseling about diet and physical activity, is recommended by professional organizations including the National Institutes of Health (NIH), Centers for Disease Control and Prevention (CDC), the American Academy of Pediatrics, and the American Academy of Family Physicians.

National surveillance of objective measures of obesity (based on measured height and weight) is limited to the National Health and Nutrition Examination Survey (NHANES), which does not provide state-, county-, or local-level estimates of obesity. Some states have worked through health care providers for surveillance purposes, while other states and cities have mandated screening of students for obesity or assessment of weight status based on BMI percentiles. In New York State, managed care programs have been required since 2005 to report as a quality indicator the percentage of adolescents who were screened for obesity (based on sex-specific BMI-for-age percentiles), counseled about nutrition, and counseled about physical activity or exercise as part of their annual preventive visit. Beginning in 2006, these data were publicly reported. In 2006, screening for obesity among adolescent patients during preventive visits was 29, 26, and 23 percent for patients enrolled in Commercial, Medicaid, and Child Health Plus managed care programs, respectively (New York State Department of Health, 2006). By 2008, these percentages were 53, 61, and 59 percent, respectively (New York State Department of Health, 2008a,b,c). Other states or health plans could expand and/or further define those criteria to include more specific or detailed counseling. In addition, insurance coverage could be expanded to cover behavioral and nutritional counseling. Whether increased screening for obesity and increased preventive counseling about diet and physical activity will lead to reduced obesity is not yet known.

Multilevel, Multisector Obesity Prevention Approaches

Repeated calls for multilevel, multisector efforts to prevent obesity have been sounded by governments, scientific and professional societies, advocacy groups, and funding agencies at the local, state, and national and international levels (IOM, 2005; Sallis et al., 2006; Story et al., 2008). Because multilevel prevention approaches are less targeted than treatment practices and reach a broader audience, they are often seen as more equitable than single-level approaches (IOM, 2005). Most obesity treatment programs are limited in both reach and long-term success. For example, while obesity prevalence, and thus the need for treatment, is higher among low-income and ethnic minority children (see Chapter 1), these populations often have inadequate access to preventive and clinical health care services (Kumanyika et al., 2008). Furthermore, ethnic minority populations tend to be underrepresented in the obesity treatment literature, reflecting their limited access to (because of both location and eligibility requirements) and/or participation in research studies (Kumanyika, 2008).

Like tobacco control initiatives, multilevel, multisector approaches combining programmatic and policy interventions with strategic, culturally appropriate, high-impact messages can change the social norms around eating, physical activity, and lifestyle behaviors. But unlike tobacco control, obesity prevention does not have a single, easily defined "villain" devoid of redemptive qualities. Food is complex—it has associations with cultural and religious traditions and provides emotional comfort, essential nutrients, and sustenance. The art of food preparation, its presentation, and feeding behaviors can also be nurturing. Food cannot be dichotomized into good and bad. It is truly essential—just not in the quantity and proportions consumed by many people in the United States and other countries where obesity is prevalent. Obesity prevention approaches that combine strategies and actions at multiple levels of the spectrum of prevention are essential; the need for such approaches and their targeted behaviors and settings are discussed below.

The Need for a Multilevel Approach

Two essential tenets of the Dietary Guidelines for Americans, which are based on the best available evidence (and when necessary, expert opinion) and form the nation's blueprint for a healthful diet, are that nutritional needs should be met through foods and beverages and that energy consumed must balance energy expended. Reducing the consumption of excess discretionary calories is a strategy that is both evidence-based and supported by expert opinion (USDA/HHS, 2005). Four of the six principal target areas identified by CDC to prevent and control obesity address specific dietary behaviors: decrease the consumption of sugar-sweetened beverages; reduce the consumption of high-calorie, energy-dense foods; increase the consumption of fruits and vegetables; and increase the initiation, duration, and exclusivity of breastfeeding.[2] For

[2] Available online: http://www.cdc.gov/obesity/stateprograms/index.html (accessed September 3, 2009).

each behavioral target, the evidence base supporting its impact on caloric intake and/or obesity differs (Khan et al., 2009).

Because many factors influence these food and beverage consumption behaviors, an intervention in a single sector or setting will be of limited value. Target behaviors are affected by federal, state, and local policies, regulations, and/or legislation; organizational practices and policies; educational campaigns; and individual knowledge and skills. For example, the significant profit factor associated with sugar-sweetened beverages and highly processed high-calorie, energy-dense foods greatly impacts the promotion of these products. A large number of factors that may mediate, enable, or amplify the intake of sugar-sweetened beverages, highly processed high-calorie, energy-dense foods, and fruits and vegetables act at multiple levels and in a variety of sectors and settings. For example, food is accessible throughout the day in multiple settings; more processed foods are available in large portion sizes and at relatively low prices; an increasing number of calories are being consumed from foods away from home; energy-dense, nutrient-poor foods are readily available in the school food environment and in the fast food and convenience store outlets in low-income urban areas; and food marketing to children has increased over the past 30 years (IOM, 2005, 2006).

Example of a Multilevel Approach: Childhood Obesity Prevention in Texas

Table 2-1 shows the efforts of an array of individuals, organizations, lawmakers, public officials, community leaders, and organizations to develop, implement, and evaluate strategies to prevent and reverse childhood obesity in Texas. A summary of a workshop convened by the IOM explains these efforts in more detail (IOM, 2009c). Texas has one of the highest obesity rates in the nation. It is home to three of the five cities with the highest obesity rates nationwide, and in 2007 two-thirds of its adult population was obese. The state's efforts target youth, with the aim of instilling healthful behaviors and lifestyles to last a lifetime. A guiding principle is that these efforts should be "evidence based, community specific, sustainable, cost-effective, and supported by effective partnerships." A goal is for the "responsibility to be broadly shared by individuals, families, communities, and the public and private sectors" (IOM, 2009c, p. 2).

The state's notable achievements have included mandating nutrition standards for school meals and vending machines in schools and enacting three bills (championed by Texas State Senator Jane Nelson) that reinstated and mandated minimum physical education requirements for elementary and middle school students in public schools and required physical fitness testing of youth in grades 3-12. It is worth noting how this was accomplished, which illustrates the importance of jurisdictional oversight in accomplishing such goals. After proposed legislation to implement school nutrition standards was unsuccessful in 2003, jurisdictional control of the nutritional environment in the public schools was shifted from the Texas Education Agency to the Texas Department of Agriculture (under the leadership of Susan Combs). This allowed the state to issue the Texas Public School Nutrition Policy, which banned foods of

TABLE 2-1 Multilevel, Multisector Approach to Childhood Obesity Prevention in Texas

Theme	Strategy	Rationale	Examples
Garnering support			
• Buy-in from the public	Focus on children and schools	Engage community members	Business leaders, parents, educators on committees
• Buy-in from legislators	Public education	Legislation, Senate Bills 19, 42, 530; Texas public school nutrition policy	Minimum physical education standards, fitness testing; nutrition standards for meals, vending machines
Modeling healthful behaviors in the community	Extend efforts beyond schools to worksites	Improve the fiscal bottom line; reduce health care costs	Workplace wellness programs; state employees
Leveraging the power of partnerships	Collaborative efforts; community-based solutions	Generate greater momentum, resources, and influence	Elected officials, state agencies, schools, worksites, higher education, urban planners, parents
Implementing comprehensive approaches	Multilevel and multisector: schools, communities, media, treatment, law enforcement	More effective than single-level interventions	Modeled after Texas Tobacco Prevention Initiative
Using social marketing	Think and act like marketers	Promote positive behaviors	Promote a healthful lifestyle, not obesity prevention
Identifying a champion	Advance a specific cause (local level), lead by example (state level)	Energize others, spread enthusiasm, communicate, unite people toward a common goal, leverage resources	Implement a program or project, pass legislation, obtain funding, incentivize change
Recognizing the power of community data	Collect local, community data; culturally adapt; collect state data	All politics are local; mobilize a call to action among decision makers	Measure outcome of interventions, compare local and state performance
Implementing federal policy change	Advocate federal policy change	Develop a foundation for states to build upon; create support for change at the community and state levels	Support federal school nutrition, physical education and fitness standards
Securing funding	Needed to launch and sustain programs; innovation important, often overlooked	Projects should be scalable, adaptable, and sustainable; funders tend to be risk-averse	Government (federal, state, local), private foundations, universities, colleges, business community

SOURCE: IOM, 2009c.

minimal nutritional value in elementary schools. After the laws requiring physical education in schools were passed, not all schools implemented physical education programs, illustrating the importance of ensuring that such laws are implemented and enforced. Nelson and Combs worked to gain public support and to gather and present data that would make a compelling case for the need for efforts to prevent obesity. When Combs became Texas State Comptroller, she was able to leverage her role in controlling the state finances to institute a grant program that covers the costs of providing physical education for middle schools attended by at least 75 percent educationally and economically disadvantaged students. It should also be noted that researchers helped develop, implement, and evaluate the Coordinated Approach to Child Health program, which has been implemented in 2,500 schools in the state. The physical fitness test that is used, the Fitnessgram, was also developed in Texas (by the Cooper Institute).

While Texas has made significant progress in changing state policies, improving the school nutrition environment, increasing students' opportunities for regular high-quality physical education, and putting evaluation systems in place, the work is not done. The state will be guided in developing best practices for the prevention and control of obesity by ensuring that these coordinated school nutrition, physical education, and health initiatives are implemented in all Texas schools; continuing to improve nutrition education and access to healthful foods; promoting worksite wellness programs; supporting comprehensive evidence-based community initiatives; and monitoring the impact of these efforts. A starting point for policy adoption, the *Texas Obesity Policy Portfolio*, was created by the Texas State Health Services Department to provide a basis for a comprehensive approach to decision making (Texas Department of State Health Services, 2006). This report is a compilation of current and proposed evidence-based obesity policy options along with their respective levels of effectiveness (untested, promising, effective), as described by the Washington State Department of Health (2005).

Conclusion

Many approaches to obesity prevention supported by varying degrees of evidence could be and are being implemented in multiple settings, cities, local communities, and states. Yet identifying where, how, and what is being done is challenging. Evaluating the process, implementation, and/or outcome of an approach is even more difficult and often is not being done at all. To address this gap, CDC has conducted a systematic review of the published literature and sought expert opinion to identify and recommend means of achieving some uniformity in approach and outcome measures (Khan et al., 2009). The results of this effort will potentially allow pooling of findings (see Chapter 8), making it easier to assess the implementation of an approach and its effectiveness and public health impact. Finally, obesity prevention initiatives need to be viewed in a broader health context and should encompass long-term, multigenerational goals, and obesity prevention policies need to be applied in a broad range of settings to achieve long-term movement toward better health.

OBESITY PREVENTION STRATEGIES IN THE BROADER CONTEXT OF OTHER PUBLIC HEALTH INITIATIVES

Strategies used to address other public health epidemics and crises, including the significant progress made in tobacco control, may be applicable in addressing the obesity epidemic (see Table 2-2) (Economos et al., 2001; Mercer et al., 2003). When the Report of the U.S. Surgeon General on smoking and health (Public Health Service, 1964) first concluded that cigarette smoking was a health hazard of sufficient importance to warrant appropriate remedial action, the evidence base regarding which strategies would be most cost-effective and/or efficacious was limited. Some 30 years later, the most promising approaches and strategies were collected and published as *Best Practices for Comprehensive Tobacco Control Programs* (CDC, 1999), later revised to include the accumulated evidence and strengthened science base (CDC, 2007). According to that report, the most successful efforts combined (1) comprehensive

TABLE 2-2 Strategies to Address Other Public Health Problems: Lessons for Obesity Prevention

Strategy	Other Public Health Problems	Obesity Prevention
Promote health (media)	• Tobacco media campaign • AIDS awareness campaigns • Drunk driving—designated driver campaigns	• *Pouring on the Pounds*—media campaign to reduce consumption of sugar-sweetened beverages in New York City (New York City Department of Health and Mental Hygiene, 2009)
Provide information	• Tobacco package warning label • Alcohol warning label	• Post calorie information on menus
Use pricing	• Tobacco tax • Alcohol tax • AIDS—free condoms, free needle exchange	• Tax sugar-sweetened beverages • Tax snacks • Subsidize fruits and vegetables • Remove subsidies for sugar-sweetened beverages and infant formula
Limit advertising	• Ban tobacco ads on television, radio	• Restrict food advertising to children and youth
Restrict access to products	• Tobacco and alcohol—restrict sales to minors • Limit placement of products to behind the counter	• Restrict sale of high-calorie, low-nutrient foods and beverages in schools
Protect public from second-hand exposure	• Tobacco—clean indoor air legislation for schools, restaurants, bars, worksites • Alcohol—ban stores near schools	• Restrict fast food outlets near schools • Offer candy-free checkout lines in grocery stores • Restrict candy and sugar-sweetened beverages in schools
Require insurance coverage for treatment or prevention	• Tobacco—cessation treatment programs, nicotine replacement therapy, quit lines	• Obesity prevention: nutrition and physical activity counseling • Obesity treatment: behavioral weight loss counseling
Frame as public health problem; change social norms and expectations	• Change social norms – Smoking restrictions in public places – Laws requiring seat belt and child car seat use – Designated driver campaigns	• Incorporate routine physical activity breaks into the workday • Provide fresh fruits and vegetables to school children in classrooms

and coordinated programmatic and policy interventions designed to influence social norms, systems, and networks; and (2) effective, sustained, and adequately funded health communication campaigns that delivered strategic, culturally appropriate, high-impact messages (CDC, 2007). Policy changes to reduce tobacco use included increasing the unit price (taxation), protecting the public from tobacco smoke (clean indoor air legislation), limiting minors' access to tobacco products, and providing insurance coverage for tobacco cessation/treatment programs (CDC, 2007). Because it is presumed that sustainable changes occur through shifts in the social environment, the report advocates interventions that influence the societal organizations, systems, and networks that facilitate desired behavior changes.

Other public health initiatives and social changes that require intervention targeting diverse groups at multiple levels may be analogous to obesity prevention (Economos et al., 2001).[3] Lessons from a range of successful attempts at social change, including increasing breastfeeding rates, seat belt use, and recycling, reveal similar keys to success. These include recognition that there was a crisis; major economic implications associated with the crisis; a science base including research, data, and evidence; leaders who could work for their cause through their knowledge, competence, talents, skills, and even cunning and charisma; coalitions to move the agenda forward and a strategic, integrated media advocacy campaign; involvement of the government at the state level to write and implement legislation and at the local level to implement change; mass communication that included consistent positive messages supported by scientific consensus and was repeated in a variety of venues; policy and environmental changes that promoted healthful behaviors; and a plan that included many components that worked synergistically (Economos et al., 2001).

Such approaches characterize the multilevel, multisector obesity prevention strategies that have been called for by governments, scientific and professional societies, advocacy groups, and funding agencies at the local, state, and national and international levels. Because many factors influence obesity-related behaviors, such strategies are believed to offer the best chances for success in obesity prevention.

REFERENCES

Brownson, R. C., D. Haire-Joshu, and D. A. Luke. 2006. Shaping the context of health: A review of environmental and policy approaches in the prevention of chronic diseases. *Annual Review of Public Health* 27:341-370.

CDC (Centers for Disease Control and Prevention). 1999. *Best practices for comprehensive tobacco control programs—April 1999*. Atlanta, GA: U.S. Department of Health and Human Services, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health.

CDC. 2007. *Best practices for comprehensive tobacco control programs—October 2007*. Atlanta, GA: Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health.

[3] Available online at http://www.preventioninstitute.org/tool_spectrum.html (accessed September 3, 2009).

CDC. 2009. *Division of nutrition, physical activity, and obesity.* http://www.cdc.gov/nccdphp/dnpao/aboutus/index.html (accessed November 30, 2009).

Cohen, L., and S. Swift. 1999. The spectrum of prevention: Developing a comprehensive approach to injury prevention. *Injury Prevention* 5(3):203-207.

Economos, C. D., R. C. Brownson, M. A. DeAngelis, S. B. Foerster, C. Tucker Foreman, J. Gregson, S. K. Kumanyika, and R. R. Pate. 2001. What lessons have been learned from other attempts to guide social change? *Nutrition Reviews* 59(3, Supplement 2):S40-S46.

Foerster, S. B., K. W. Kizer, L. K. DiSogra, D. G. Bal, B. F. Krieg, and K. L. Bunch. 1995. California's '5 a day—for better health!' campaign: An innovative population-based effort to effect large-scale dietary change. *American Journal of Preventive Medicine* 11(2):124-131.

HHS (U.S. Department of Health and Human Services). 2008. *2008 physical activity guidelines for Americans.* Office of Disease Prevention and Health Promotion Publication No. U0036. Washington, DC: U.S. Department of Health and Human Services.

IOM (Institute of Medicine). 1995. *Weighing the options: Criteria for evaluating weight-management programs.* Edited by P. R. Thomas. Washington, DC: National Academy Press.

IOM. 2005. *Preventing childhood obesity: Health in the balance.* Edited by J. Koplan, C. T. Liverman, and V. I. Kraak. Washington, DC: The National Academies Press.

IOM. 2006. *Food marketing to children and youth: Threat or opportunity?* Edited by J. M. McGinnis, J. Gootman and V. I. Kraak. Washington, DC: The National Academies Press.

IOM. 2007. *Progress in preventing childhood obesity: How do we measure up?* Edited by J. Koplan, C. T. Liverman, V. I. Kraak, and S. L. Wisham. Washington, DC: The National Academies Press.

IOM. 2009a. *Local government actions to prevent childhood obesity.* Edited by L. Parker, A. C. Burns, and E. Sanchez. Washington, DC: The National Academies Press.

IOM. 2009b. *Community perspectives on obesity prevention in children: Workshop summaries.* Edited by P. T. Whitacre and A. C. Burns. Washington, DC: The National Academies Press.

IOM. 2009c. *Childhood obesity prevention in Texas: Workshop summary.* Edited by K. Nyberg, A. C. Burns and L. Parker. Washington, DC: The National Academies Press.

Khan, L. K., K. Sobush, D. Keener, K. Goodman, A. Lowry, J. Kakietek, and S. Zaro. 2009. Recommended community strategies and measurements to prevent obesity in the United States. *Morbidity and Mortality Weekly Report* 58(RR-7):1-26.

Kumanyika, S. 2008. Ethnic minorities and weight control research priorities: Where are we now and where do we need to be? *Preventive Medicine* 47(6):583-586.

Kumanyika, S., R. W. Jeffery, A. Morabia, C. Ritenbaugh, and V. J. Antipatis. 2002. Obesity prevention: The case for action. *International Journal of Obesity* 26(3):425-436.

Kumanyika, S. K., E. Obarzanek, N. Stettler, R. Bell, A. E. Field, S. P. Fortmann, B. A. Franklin, M. W. Gillman, C. E. Lewis, W. C. Poston II, J. Stevens, and Y. Hong. 2008. Population-based prevention of obesity: The need for comprehensive promotion of healthful eating, physical activity, and energy balance: A scientific statement from American Heart Association Council on Epidemiology and Prevention, Interdisciplinary Committee for Prevention (formerly the Expert Panel on Population and Prevention Science). *Circulation* 118(4):428-464.

Larson, N. I., M. T. Story, and M. C. Nelson. 2009. Neighborhood environments: Disparities in access to healthy foods in the U.S. *American Journal of Preventive Medicine* 36(1):74-81.e10.

Lovasi, G. S., M. A. Hutson, M. Guerra, and K. M. Neckerman. 2009. Built environments and obesity in disadvantaged populations. *Epidemiologic Reviews* 31:7-20.

McLeroy, K. R., D. Bibeau, A. Steckler, and K. Glanz. 1988. An ecological perspective on health promotion programs. *Health Education Quarterly* 15(4):351-377.

Mello, M. M. 2009. New York City's war on fat. *New England Journal of Medicine* 360(19):2015-2020.

Mercer, S. L., L. W. Green, A. C. Rosenthal, C. G. Husten, L. K. Khan, and W. H. Dietz. 2003. Possible lessons from the tobacco experience for obesity control. *The American Journal of Clinical Nutrition* 77(4, Supplement):1073S-1082S.

Moore, L. V., A. V. Diez Roux, J. A. Nettleton, and D. R. Jacobs, Jr. 2008. Associations of the local food environment with diet quality—a comparison of assessments based on surveys and geographic information systems: The multi-ethnic study of atherosclerosis. *American Journal of Epidemiology* 167(8):917-924.

Morland, K., S. Wing, and A. Diez Roux. 2002. The contextual effect of the local food environment on residents' diets: The atherosclerosis risk in communities study. *American Journal of Public Health* 92(11):1761-1767.

New York City Department of Health and Mental Hygiene. 2009. *New campaign asks New Yorkers if they're "pouring on the pounds"*. http://www.nyc.gov/html/doh/html/pr2009/pr057-09.shtml (accessed March 8, 2010).

New York State Department of Health. 2006. *2006 New York State managed care plan performance: A report on quality, access to care, and customer satisfaction.* http://www.health.state.ny.us/health_care/managed_care/qarrfull/qarr_2006/qarr2006.pdf (accessed April 5, 2010).

New York State Department of Health. 2008a. *Adolescent preventive care—statewide—commercial managed care.* http://www.health.state.ny.us/health_care/managed_care/reports/eqarr/2008/statewide/commercial/adolescent_preventive_care.htm (accessed April 6, 2010).

New York State Department of Health. 2008b. *Adolescent preventive care—statewide—Medicaid managed care.* http://www.health.state.ny.us/health_care/managed_care/reports/eqarr/2008/statewide/medicaid/adolescent_preventive_care.htm (accessed April 6, 2010).

New York State Department of Health. 2008c. *Adolescent preventive care—statewide—child health plus managed care.* http://www.health.state.ny.us/health_care/managed_care/reports/eqarr/2008/statewide/child_health_plus/adolescent_preventive_care.htm (accessed April 6, 2010).

Public Health Service. 1964. *Smoking and health: Report of the Advisory Committee to the Surgeon General of the Public Health Service.* Public Health Service Publication No. 1103. Washington, DC: U.S. Department of Health, Education, and Welfare.

Public Health Service. 2001. *The Surgeon General's call to action to prevent and decrease overweight and obesity.* Rockville, MD: U.S. Department of Health and Human Services.

Rattray, T., W. Brunner, and J. Freestone. 2002. *The new spectrum of prevention: A model for public health practice.* Martinez, CA: Contra Costa Health Services.

Sallis, J. F., R. B. Cervero, W. Ascher, K. A. Henderson, M. K. Kraft, and J. Kerr. 2006. An ecological approach to creating active living communities. *Annual Review of Public Health* 27:297-322.

Stokols, D., J. Allen, and R. L. Bellingham. 1996. The social ecology of health promotion: Implications for research and practice. *American Journal of Health Promotion* 10(4):247-251.

Story, M., K. M. Kaphingst, R. Robinson-O'Brien, and K. Glanz. 2008. Creating healthy food and eating environments: Policy and environmental approaches. *Annual Review of Public Health* 29:253-272.

Texas Department of State Health Services. 2006. *Texas obesity policy portfolio*. Austin, TX: Center for Policy and Innovation.

USDA (U.S. Department of Agriculture)/HHS. 2005. *Dietary guidelines for Americans, 6th ed*. Home and Garden Bulletin No. 232. Washington, DC: U.S. Government Printing Office.

Washington State Department of Health. 2005. *Nutrition and physical activity: A policy resource guide*. http://www.dshs.state.tx.us/obesity/pdf/WAnutritionphysicalactivityguide.pdf (accessed December 1, 2009).

WHO (World Health Organization). 2000. *Obesity: Preventing and managing the global epidemic. Report of a WHO consultation*. WHO Technical Report Series 864. Geneva: World Health Organization. http://www.who.int/nutrition/publications/obesity/en/index.html.

WHO. 2003. *Diet, nutrition, and the prevention of chronic diseases. Report of a joint WHO/FAO expert consultation*. WHO Technical Report Series 916. Geneva: World Health Organization.

WHO. 2004. *Global strategy on diet, physical activity, and health*. Geneva: World Health Organization.

Wong, F., M. Huhman, C. Heitzler, L. Asbury, R. Bretthauer-Mueller, S. McCarthy, and P. Londe. 2004. VERB™—a social marketing campaign to increase physical activity among youth. *Preventing Chronic Disease* 1(3):1-7.

World Cancer Research Fund and American Institute for Cancer Research. 2007. *Food, nutrition, physical activity, and the prevention of cancer: a global perspective*. Washington, DC: American Institute for Cancer Research.

Zenk, S. N., L. L. Lachance, A. J. Schulz, G. Mentz, S. Kannan, and W. Ridella. 2009. Neighborhood retail food environment and fruit and vegetable intake in a multiethnic urban population. *American Journal of Health Promotion* 23(4):255-264.

3

Rationale for and Overview of the L.E.A.D. Framework

KEY MESSAGES

- The current research literature lacks the power to set a clear direction for obesity prevention interventions across a range of target populations.
- The concepts of evidence-based decision making apply to medicine and to numerous disciplines and approaches in addition to public health, social work, and public policy, with common principles and themes.
- To apply the principles of evidence-based public health or public policy to obesity prevention requires an expanded approach to locating, evaluating, and assembling evidence that is oriented toward meeting the needs of decision makers. Such an approach can be guided by the different types of questions decision makers need to answer.
- The L.E.A.D. framework (for **L**ocate Evidence, **E**valuate Evidence, **A**ssemble Evidence, Inform **D**ecisions) introduced in this chapter is designed to (1) assist policy makers and others who make decisions related to obesity prevention in identifying and evaluating the best evidence available and summarizing it for use, (2) assist various audiences in packaging the evidence in the most appropriate way, (3) motivate and guide researchers in multiple disciplines and settings to generate high-quality evidence with strong program and policy relevance to obesity prevention, and (4) improve approaches to the use of evidence to inform decisions about complex population health problems in general.

Ideally there would be a firm and comprehensive evidence base to inform the myriad decision makers whose actions—intentionally or unintentionally—influence the social, policy, and environmental determinants of excess weight gain and the resulting overweight/obesity. In reality, however, the concept of an evidence base as applied to decision making about treatments for established obesity or other clinical problems does not apply to many of the decisions that must be made to prevent obesity. The current research literature lacks the power to set a clear direction for obesity prevention interventions across a range of target populations. Thus, although the concept of evidence-based obesity prevention is likely to resonate with many health professionals, policy makers, and stakeholders, challenges are associated with the identification of data that can be used to address the different types of questions decision makers might need to answer. A range of evidence-based intervention approaches have been developed for such public health issues as tobacco control and immunization. However, there is often a mismatch between the types of evidence that are available to answer public health questions and the types of evidence being generated and published.

With respect to obesity prevention, research into the causal factors driving dietary or energy intake and physical activity levels includes a broad range of designs and evidence, including randomized controlled (experimental) trials and quasi-experimental studies at the individual, organizational, and community levels. Although a similarly broad array of research approaches is potentially applicable to environmental and policy strategies to reduce obesity risk through improved diet and reduced caloric intake and/or increased physical activity, the basic methodologies require different conceptualizations, implementation, and emphases for obesity prevention research—particularly to inform interventions designed to change policies or environmental contexts for eating and physical activity. Although the literature reports some positive findings with respect to strategies that work for obesity prevention, there are many gaps in the evidence. This chapter describes those gaps and reviews concepts from evidence-based public health and public policy that can provide the fundamentals for the use of evidence to address complex population health problems such as obesity prevention. The focus is on the evidence needs most likely to impede progress in obesity prevention and those considered most important to address in the development of an evidence framework to inform decisions about obesity prevention interventions.

STATUS OF THE EVIDENCE BASE

The contrast between the high prevalence and consequent importance of addressing obesity and the paucity of the knowledge base with which to inform prevention efforts is striking. This evidence gap creates tension between the sense of urgency to take action and the lack of specificity about what actions to undertake. This is especially true in circumstances when decision makers cannot wait for a body of literature to be created and critically reviewed.

The evidence for obesity prevention is ideally accumulated from a variety of sources to provide insight into a particular topic, often combining quantitative and qualitative data. To this end, systematic reviews of the intervention literature are undertaken that apply strategies designed to limit bias in the assembly, critical appraisal, and synthesis of all relevant studies on a specific topic (Cook et al., 1995). As noted in Chapter 1, the challenges in evaluating and assembling evidence related to obesity prevention have been highlighted in previous Institute of Medicine (IOM) reports (IOM, 2005, 2007).

To obtain an overview of the status of the existing relevant evidence base, the committee examined studies published over the past 13 years to identify reviews focused on obesity prevention. This examination was not meant to be exhaustive, but to illustrate the scope of existing appraisals and the criteria used to determine what studies did or did not qualify for inclusion. Nearly 50 such reviews, hereafter referred to as “appraisals,” were identified (see Appendix C). Approximately half of these appraisals consider literature published as early as 1980, and several include literature from as early as 1966 and/or the inception of the search databases used. The appraisals consist of meta-analyses, systematic reviews, integrative reviews, a review of reviews, evidence syntheses, best practice summaries, and task force recommendations. Nearly one-third of the appraisals were published in journals specific to obesity; the other two-thirds appeared in journals focused on medicine, preventive medicine, public health, health promotion, family and community health, health education, nutrition, nursing, epidemiology, endocrinology and metabolism, and psychology. Almost three-quarters of the appraisals had been published since 2005, indicating the increased interest in the topic. This examination of the literature revealed the challenges involved in applying traditional evidence hierarchies to population-based prevention efforts, as described below.

Quantity of Available Evidence

In the appraisals examined, the number of studies found eligible for inclusion ranges from 3 to 158, a small percentage of the number of studies initially identified for potential further analysis, which ranged from 12 to 13,158. However, while all of the appraisals report the number of studies found eligible for inclusion, some do not report the number of studies initially identified. The range of studies included reveals a clear discrepancy between what authors of specific studies might think of as obesity prevention research and what authors of reviews consider to be part of the evidence base. There is considerable overlap in studies included in these appraisals; however, conclusions vary as to what is effective. Furthermore, most studies included in the appraisals lack detail about the process of the interventions and process evaluation information. Several of the appraisals explicitly address what types of evidence should be considered relevant, whether one fails to be comprehensive by restricting evidence to randomized controlled trials (RCTs) with rigorous inclusion and exclusion criteria, and whether literature searches should extend beyond well-known databases and include monographs or reports not published in peer-reviewed journals.

Lack of Conceptual Frameworks

The lack of a conceptual framework specifically related to evidence selection has resulted in the majority of appraisals failing to draw conclusions about the most effective components of interventions. Many of the appraisals have a narrow focus (e.g., schools, local environments) and fail to recognize the larger systems context in which the interventions were implemented. This lack of focus on contextual considerations (such as feasibility [including legal and political aspects], sustainability, effects on equity, potential side effects, acceptability to stakeholders, and cost-effectiveness) and other dimensions of generalizability (such as the intervention reach and the representiveness of participants, implementation and adaptation of the intervention, and its maintenance and institutionalization) limits the usefulness of the appraisals to inform decision making in diverse settings. Furthermore, the challenges that arise when interventions are undertaken in the real world lead the conclusions drawn to be negative or obsolete. In fact, many appraisals exclude comprehensive multilevel, multisector interventions (the most promising type of interventions for addressing obesity, as discussed in Chapter 2) because the specific effects on individuals cannot be disaggregated, and the results are considered "messy." When such interventions target multiple behaviors simultaneously, it is often difficult to disentangle their impacts. Although such trials yield some positive results, the many gaps in the evidence suggest that obesity prevention lacks the vast array of effective interventions that are available for many other public health issues (e.g., tobacco control, immunization). While efforts are under way to fill these gaps, the gravity of the obesity epidemic calls for immediate action, suggesting that, as noted in an earlier IOM report, decision makers should rely on the best available evidence, instead of waiting for the best possible evidence (IOM, 2005).

Choice of Outcomes

The choice of outcomes on which to focus affects conclusions about the effectiveness of interventions across this set of appraisals. The outcome measures for selecting which individual studies to include in the appraisals were similar, and in most cases were cause for excluding a large percentage of studies initially identified for consideration. As stated in Chapter 2, the ultimate objectives of obesity prevention efforts are lowering the mean BMI level and decreasing the rate at which people enter the upper end of the BMI distribution. Accordingly, most individual-level studies focus on distal measures of weight status (e.g., percent overweight or obese, BMI or BMI z-score, body fat percentage), risk factors, or disease presence rather than the more proximal psychosocial and behavioral outcomes and the pathways that can alter the conditions in which eating, physical activity, and weight control occur (see Figure 2-2 in Chapter 2). Moreover, the period of follow-up in many of the studies included in the appraisals may not have been long enough to detect significant, sustainable changes in weight status. Indeed, the comprehensive changes called for to combat the obesity epidemic—changes in the physical environment, social norms, cultural practices, and policy—often require years to accomplish and are not always tracked or funded at adequate levels for sustained periods. In the appraisals that specify duration of study or length of follow-up as an inclusion criterion, the vast majority required a period of

6-12 months. Thus, many of the appraisals examined report null or negative findings, which limits the ability to justify scaling up of promising interventions. Yet while the ultimate biological outcome (changes in weight status) may not have been achieved, intermediate outcomes (structural, institutional, systemic, social, and environmental changes in society or the community, as shown in Figure 2-2) that influence the cognitive, social, and behavioral outcomes among individuals may have been achieved. These intermediate outcomes have the potential to lead to biological outcomes if the interventions are sustained.

Specifically for evidence-based obesity prevention efforts, a body of intervention research on policy and environmental approaches is largely absent from the literature. Some recent studies examine policy- and environmental-level strategies with the potential to reduce obesity. These studies consider how changes in the food and physical activity environments influence eating and activity behaviors. The criteria applied to rate these efforts include reach, mutability, transferability, effect size, and sustainability of health impact, which may be useful in informing scalability.

FUNDAMENTAL EVIDENCE CONCEPTS

The Perspective from Evidence-Based Medicine

The movement toward evidence-based practice in various public service arenas, including obesity prevention, has been significantly influenced by earlier trends in medicine and published guidelines for evidence standards in evidence-based medicine. For example, the Cochrane Collaboration—an international, independent nonprofit organization dedicated to making current, accurate information about the effects of health care available—assembles, maintains, and disseminates systematically collected and reviewed information on health care interventions (The Cochrane Collaboration, 2009). Its major product is the Cochrane Collaboration of Systematic Reviews (published along with several other databases in the Cochrane Library), which focuses solely on synthesizing research-based evidence on the effectiveness of various treatments in specific circumstances. The Cochrane Library is intended to serve as an electronic resource and reference guide for physicians who make individual patient care decisions. It should be noted that the Cochrane Collaboration has broadened its original scope and now includes the Cochrane Public Health Review Group, which produces and publishes Cochrane reviews of the effects of population-level public health interventions.

The criteria these databases use to select studies from the medical literature to synthesize focus heavily on methodological standards. The level of certainty (internal validity) is a paramount criterion in screening for high-quality evidence. RCTs, which involve randomization of subjects in a double-blinded design, are viewed as one of the best means of maximizing the level of certainty of the causal relationship being tested, affording the tightest possible controls for testing the effectiveness of a treatment and providing the strongest basis for drawing causal inferences about interventions (West et al., 2000).

The evidence selection criteria of evidence-based medicine have led to a widely disseminated evidence hierarchy that has been adopted by evidence-based movements in nursing, psychology, social work, education, and public health. The hierarchy characterizes the quality of evidence for decision making at three levels. Systematic reviews of RCTs and individual RCT studies of particular therapies are at the top of this hierarchy, serving as the "gold standard"; next are observational studies and systematic reviews of such studies; and at the lowest level are physiologic studies and unsystematic clinical observations (Guyatt and Rennie, 2002). Other research designs and evidence sources, such as surveys, quasi-experimental designs, and qualitative research methods, are sometimes excluded from the evidence hierarchy of evidence-based medicine.

Leading authors in evidence-based medicine have acknowledged that the evidence hierarchy applies better to medical therapies than to clinical decisions that lie outside the therapy domain (Guyatt and Rennie, 2002). For example, evidence on issues of diagnosis, prognosis, or assessment of harmful side effects of a therapy (where an "intervention" cannot be randomly assigned or manipulated) call for alternative study designs and evidence selection criteria.

Evidence-Based Decision Making in Public Health

Many of the approaches discussed in this report fall within the broad category of evidence-based public health (EBPH), which has been defined as "the process of integrating science-based interventions with community preferences to improve the health of populations" (Kohatsu et al., 2004, p. 419). Formal discourse on the nature and scope of EBPH originated in the 1990s. In 1997, Jenicek first defined the concept (Jenicek, 1997). In 1999, scholars and practitioners in Australia (Glasziou and Longbottom, 1999) and the United States (Brownson et al., 1999) elaborated further on EBPH. Glasziou and Longbottom (1999) posed a series of questions to enhance its uptake (e.g., "Does this intervention help alleviate this problem?") and identified 14 sources of high-quality evidence. Brownson and colleagues (1999, 2003) described a six-stage process by which practitioners can take a more evidence-based approach to decision making. In 2004, Rychetnik and colleagues developed a glossary summarizing many key aspects of EBPH (Rychetnik et al., 2004).

Following principles of EBPH when addressing obesity prevention has numerous direct and indirect benefits, including access to more and higher-quality information on what works, a greater likelihood of implementing successful programs and policies, and more efficient use of public and private resources (Brownson et al., 2003, 2009a; Cavill et al., 2006; Hausman, 2002; Kohatsu and Melton, 2000). As these principles have been formalized, two parallel developments have occurred: recognition that science needs to play a key role in the development of public health programs and policies, and the instant availability of greater amounts of information (Anderson et al., 2005). The principles of evidence-based decision making apply to medicine and to numerous disciplines and approaches in addition to public health, social work, and public policy (the latter being referred to as evidence-based public policy [EBPP]) (Brownson et al., 2009b; Choi, 2005; Dobrow et al., 2004; Kerner, 2008; Satterfield et al., 2009).

Both EBPH and EBPP appear to be informed by consensus that a combination of scientific evidence and values, resources, and context should enter into decision making (Brownson et al., 2003; Muir Gray, 2009; Rychetnik et al., 2004; Satterfield et al., 2009). In EBPP, three key domains have been described: (1) *process*, which involves an attempt to understand approaches to enhancing the likelihood of policy adoption (e.g., a strong coalition would enhance the likelihood of advancing an agenda for obesity prevention); (2) *content*, which involves identifying specific policy elements that are likely to be effective (e.g., obesity policy should be grounded in the latest science); and (3) *outcomes*, which involves documenting the potential impact of a policy (e.g., a system should be in place to track the effects of new obesity prevention policies) (Brownson et al., 2009b). While the use of research-derived evidence may be a key feature of most policy models, scientific evidence will not necessarily carry as much weight as other types of evidence in real-world policy-making settings. Policy makers operate according to a different hierarchy of evidence than scientists (Choi et al., 2005), leaving the two groups to work in "parallel universes" (Brownson et al., 2006). In interviews with decision makers (which included state legislators and public health agency officials), many reported that they were not trained to distinguish between good and bad data and were therefore prone to the influence of misused "facts" often presented by interest groups (Jewell and Bero, 2008).

Types of Questions That Need to Be Answered

Several types of questions have been identified in the field of EBPH that are important to consider in relation to different aspects of obesity prevention (Brownson et al., 2009b). Type 1 questions address the causes of obesity-related diseases and the magnitude, severity, and preventability of obesity-related risk factors. They suggest that "something should be done" about the obesity epidemic. These questions (whether, where, when) might include such issues as the burden of disease, etiologic relationships, time trends, and high-risk populations. Type 2 questions address the relative impact of specific interventions that do or do not improve health, adding "specifically, this should be done." For example, Type 2 questions (what works, in what settings, with what outcomes, and at what costs) might include such issues as where interventions are effective, what outcomes are observed, and what an intervention costs. Type 3 questions address how and in what context interventions are implemented and how they are received, thus informing "how something should be done." These questions might address how interventions may need to be adapted, what training is needed for staff, or how to sustain effects over time.

Two terms relevant to evidence quality and standards in evidence-based medicine or public health are key to the discussion here and throughout this report: *internal validity* (the level of certainty of the causal relationship between an intervention and the observed outcomes), hereafter referred to as "level of certainty," and *external validity* (the extent to which research results can be generalized to other populations, including individuals, settings, contexts, and time frames, which speaks to issues of the transferability and application of findings to real-world settings), hereafter referred to as "generalizability." (See Appendix A for formal definitions.)

In relation to Type 2 and perhaps Type 3 questions, studies to date have tended to emphasize the level of certainty of the causal relationship between an intervention and the observed outcome (e.g., well-controlled efficacy trials) while giving limited attention to the intervention's generalizability (e.g., the translation of science to the various circumstances of practice) (Glasgow, 2008; Glasgow et al., 2006; Green and Glasgow, 2006; Klesges et al., 2008). For example, Klesges and colleagues (2008) reviewed 19 childhood obesity studies to assess the extent to which dimensions of external validity were reported. Their work reveals that some key contextual variables (e.g., cost, program sustainability) are missing entirely in the peer-reviewed literature on obesity prevention. Because generalizability is deemphasized or compromised in study selection and evidence screening procedures based on the hierarchical framework of evidence-based medicine, concern has been raised that applying that hierarchy's narrow inclusion criteria results in a limited pool of sources for subsequent evidence synthesis (Dixon-Woods et al., 2001). Further, it is increasingly acknowledged that the traditional hierarchy works less well in public health and health care arenas involving large-scale, complex social interventions than in the arena of medical therapies. Evidence standards for community-based public health problems call for taking a broader perspective to locate useful forms of evidence. The types of evidence that are available for assessing both level of certainty and generalizability were a key consideration when the committee developed the framework proposed in this report.

Approaches to Evaluating Evidence

While the highest-quality source of evidence is generally deemed to be the RCT, much of the evidence available for environmental and policy strategies to reduce obesity risk is not derived from such trials. For example, a randomized design is seldom useful in policy research because the scientist cannot randomly assign exposure (the policy), and understanding of the policy process is often gained from qualitative methods (e.g., case studies). In addition, reductionist approaches for inferring causality (common in medicine) tend to reduce a problem to a single cause (Davidovitch and Filc, 2006), whereas obesity is likely to be influenced by a complex web of causation (Hill and Peters, 1998; Kumanyika and Brownson, 2007; Kumanyika et al., 2002).

Attention to Contextual Issues That Influence Interventions

While numerous authors have written about the role of context in informing evidence-based practice and policy, there is little consensus on its definition. When moving from clinical interventions to population-level and policy interventions, context becomes more uncertain, variable, and complex (Dobrow et al., 2004). One useful definition of context highlights information needed to adapt and implement an evidence-based intervention in a particular setting or population (Rychetnik et al., 2004).

The context for questions on how to implement an obesity-related intervention involves five overlapping domains. First is characteristics, such as level of education, of the target population. Second, interpersonal variables provide important context; for example, a person who comes from a family that exercises together may have a lower risk of developing obesity. Third, organizational variables should be considered, such as how an agency's capacity (e.g., a trained workforce, agency leadership) may influence its success in carrying out an evidence-based obesity prevention program (Dreisinger et al., 2008; Hausman, 2002). Fourth, social norms and culture are known to shape many health behaviors. Finally, larger political and economic forces affect context; for example, a high rate of obesity may influence a state's political will to address the issue in a meaningful and systematic way.

Particularly for high-risk and understudied populations, there is a pressing need for evidence on obesity prevention in relation to contextual variables and ways of adapting programs and policies across settings and population subgroups. Contextual issues are being addressed more fully in the new "realist review" approach—a systematic review process that involves examining not only *whether* an intervention works but also *how* it works in real-world settings (Pawson et al., 2005). This enhanced understanding of context is a key theme of a *systems approach* (see Chapter 4). Well-reasoned recommendations have emerged from many quarters to replace the hierarchy of evidence-based medicine with a systems approach to garnering, analyzing, and appraising evidence when addressing complex public health problems such as obesity prevention (Briss et al., 2000; Brownson et al., 2003; Green and Glasgow, 2006; Harris, 2000; McKinlay, 1992; Petticrew and Roberts, 2003; Rychetnik et al., 2004; Swinburn et al., 2005). There are several reasons for this shift.

First, unlike medical therapies focused on individual patient care, obesity prevention interventions are typically mounted on a relatively large scale and focus on health issues affecting communities at large (at the regional, national, or even global level). Causal pathways to population outcomes do not involve the straightforward, patient-level treatment-to-outcome linkages typically tested in trials of medical therapies.

Second, obesity prevention interventions frequently comprise multiple components operating at different levels of society. In a typical example, a legislative action at the highest level (e.g., a state action on mandatory nutrition labeling) may be combined with health campaigns and education programs administered at the community level (e.g., a county or district program delivered via local agencies such as schools, clinics, and hospitals), which are then expected to affect factors at the household/individual level (e.g., lifestyle choices and eating behaviors). Different intervention components at different levels are mediated by social and human factors operating in their natural settings. All such factors call for appropriate observation, measurement, and modeling that are constrained by traditional experimental and quasi-experimental designs. Complex interventions often are constrained in simplistic experimental designs and accompanying analytic models that assume a treatment to be a single manipulated factor (similar to a medical therapy).

Third, while the level of certainty (or internal validity) of research designs continues to be salient, external validity increases greatly in importance when one is identifying and appraising research evidence on public health issues such as obesity. Effective public health interventions need to have generalizability, transferability, and sustainability beyond small-scale research studies (Green and Glasgow, 2006).

Finally, not all questions that arise in taking social action are best answered by cause–effect designs or even research-based evidence. While the issue of whether a particular treatment works to reduce obesity may be answered with experimental designs, questions such as what causative and protective factors could potentially be targeted by interventions call for different approaches and research methods. Correspondingly, different criteria must be applied to build the needed evidence base. In short, the type of evidence must be matched to the policy or practice objective at hand. Potential alternative research designs to the RCT are described in Appendix E.

Transparency in Decision Making

Transparency implies openness in decision making, effective communication, and accountability. The ability to make public health decisions as transparent as possible is likely to foster public trust, better risk communication, and financial accountability (Honore et al., 2007; Laforest and Orsini, 2005; McGloin et al., 2009). Few frameworks for evidence-based decision making have made transparency a central theme; the committee decided in its deliberations that transparency was a goal of its framework. Transparency in summarizing and communicating the evidence used in making a decision is discussed in Chapter 7.

Approaches to Measurement

Measurement of progress in preventing obesity can take several different forms. The recommended indicator of obesity for adults and children is BMI—calculated as weight in kilograms divided by the square of height measured in meters (kg/m^2). Beyond measurement at the individual level, consideration of how to measure the effects of environmental and policy changes is a relatively new aspect of obesity research, and appropriate measures and evaluation designs are still being developed. Chapter 8 and Appendix E provide more detailed discussion of methods for measuring effects and generating evidence that are applicable to decision making on obesity prevention and other public health issues.

For decisions on environmental and policy interventions, both quantitative data (e.g., epidemiologic) and qualitative information (e.g., analysis of the content of policy, narrative accounts of the policy process) are often important. Quantitative evidence (i.e., data in numerical quantities) can take many forms, ranging from scientific information in peer-reviewed journals, to data from public health surveillance systems, to evaluations of individual programs or policies. Qualitative evidence involves non-numerical observations, collected by such methods as participant observation, group interviews, or focus groups. Qualitative evidence can be presented in narrative form as a powerful means of influencing policy deliberations, priority setting, and the proposing of environmental and policy solutions by telling persuasive stories with an emotional hook and intuitive appeal.

In designing obesity prevention interventions aligned with population needs and profiles, data on stakeholder opinions, preferences, and acceptability could serve as a useful and legitimate part of the evidence base. Swinburn and colleagues (2005), on behalf of the Prevention Group of the International Obesity Task Force (IOTF), proposed an evidence framework for obesity prevention that incorporates a broad policy perspective to guide users in locating a pool of evidence relevant to the problem. This framework acknowledges that valid and useful evidence cannot be restricted to impact evidence from traditional RCTs. Policy making, from the IOTF perspective, should be guided by a goal to reduce uncertainty given the best information available on the issue. Applying this criterion results in the inclusion of case histories, simulation studies, expert (informed) opinion, economic analyses, formative program evaluations, monitoring and surveillance data, and logic models, as well as research-based evidence from both experimental and observational studies, as relevant data sources.

Evaluation of specific environmental and policy interventions is complicated by the fact that additive or synergistic effects of multiple interventions across different levels and sectors may be necessary to achieve an impact on behaviors related to energy balance and to have a measureable impact on weight status. This need can be addressed in part by multilevel interventions or combinations of studies, but to date such studies are few in number. While the development of metrics for environmental and policy interventions is still in an early (first-generation) stage, several sources provide useful discussion of methods for data collection, promising indicators, and future directions for the field (Brownson et al., 2009a; Cheadle et al., 2000; Khan et al., 2009; Lytle, 2009; McKinnon et al., 2009).

APPLICATION OF AN EXPANDED EVIDENCE FRAMEWORK

To apply the principles of EBPH or EBPP to obesity prevention, an expanded approach to locating, evaluating, and assembling evidence is needed. Neither the more complex nature of the evidence nor the attendant controversies can be a reason for inaction or for taking actions purely on the basis of guesswork. In addition, evidence on the effectiveness or impact of ongoing actions is needed. Accordingly, this report provides a framework for identifying and using scientific and other relevant knowledge to address complex public health problems, and advocates for its proactive use in addressing the obesity epidemic to increase the soundness of decisions being made on a daily basis by a variety of public and private actors.

Key considerations in the proposed framework include the following:

- The framework should include an understanding and characterization of the types of decisions and decision-making processes and contexts that are applicable. These decisions often take place within complex systems.
- Building on this premise requires an understanding and characterization of the respective types of questions being asked, the context involved, and evidence that can inform the decisions being made, along with algorithms to guide the use of this evidence in decision making.
- To give decision making priority as the basis for the framework does not subordinate the importance of science as a basis for action. Rather, the committee's approach emphasizes the role of science in helping to address the challenges associated with the untenable societal burden of obesity. The committee's approach also broadens the scientific platform so that the focus goes beyond obesity as a biomedical problem to encompass multiple disciplines, recognizing that obesity can be prevented only by interventions spanning the social, economic, cultural, and policy realms. This broader platform opens doors to a wealth of additional approaches for addressing evidence needs in the obesity prevention arena.

THE L.E.A.D. FRAMEWORK

The framework developed by the committee for obesity prevention decision making, and by extension for decision making about other complex prevention scenarios, reflects the above considerations (see Figure 3-1). It incorporates concepts and approaches that are standard procedure in the development of practice guidelines: formulate the question, use an analytic framework to specify key questions to be addressed, do a systematic literature search, evaluate the evidence, describe the body of evidence, and determine its implications. The committee adapted this general approach for use in policy and programmatic decision making about complex public health problems such as obesity prevention.

The **L**ocate Evidence, **E**valuate Evidence, **A**ssemble Evidence, Inform **D**ecisions (L.E.A.D.) framework, which highlights the four steps that are at the core of the evidence review process, is intended to convey the entire picture of obesity prevention decision making. It was designed to identify the nature of the evidence that is needed; to clarify what changes in current approaches to generating and evaluating evidence will help meet decision makers' evidence needs; and to facilitate a systematic approach that enables identification, implementation, and evaluation of promising, reasonable actions. The framework is intended as a guide to using evidence to inform decision

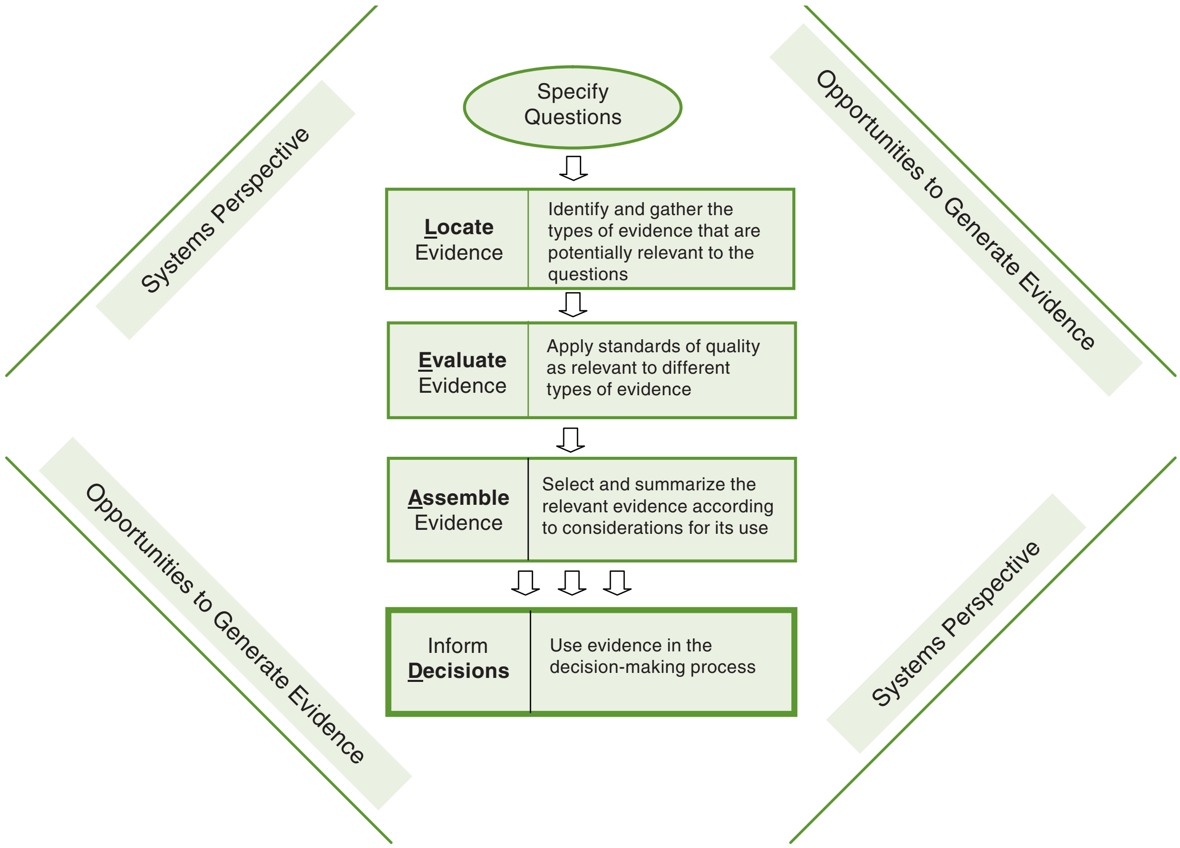


FIGURE 3-1 The **L**ocate Evidence, **E**valuate Evidence, **A**ssemble Evidence, Inform **D**ecisions (L.E.A.D.) framework for obesity prevention decision making.

making. Evidence—regardless of how complete or relevant—is only one component of making a decision; decision makers usually must take one or more other factors into account when making final decisions.

For a conceptual understanding of the framework, it is useful to begin with the systems perspective that surrounds the steps in the framework to emphasize the importance of taking such a perspective throughout the process, and then move to specifying questions that will guide the search for evidence; locating, evaluating, and assembling the evidence; and making a decision based on this evidence. Finally, opportunities to generate evidence, like a systems perspective, are pertinent throughout the process. Although the framework leads decision makers and investigators through a series of steps, it is important to point out that a user could begin at any point on the framework and return to earlier steps—for example, from the assemble step back to further elaboration of the questions, broadening the search for evidence accordingly. For example, a researcher seeking ideas for solution-oriented research could first attempt to identify user questions or examine a report generated during the decide step and work backward to identify research needs encountered in locating or evaluating evidence.

Section II of this report provides detail on each step of the L.E.A.D. framework, and elaborates on the other concepts the framework emphasizes: the importance of taking a systems perspective and opportunities to generate evidence. By applying this framework as outlined, decision making, and ultimately public health practice, can be improved.

REFERENCES

Anderson, L. M., R. C. Brownson, M. T. Fullilove, S. M. Teutsch, L. F. Novick, J. Fielding, and G. H. Land. 2005. Evidence-based public health policy and practice: Promises and limits. *American Journal of Preventive Medicine* 28(5, Supplement):226-230.

Briss, P. A., S. Zaza, M. Pappaioanou, J. Fielding, L. Wright-De Aguero, B. I. Truman, D. P. Hopkins, P. D. Mullen, R. S. Thompson, S. H. Woolf, V. G. Carande-Kulis, L. Anderson, A. R. Hinman, D. V. McQueen, S. M. Teutsch, and J. R. Harris. 2000. Developing an evidence-based Guide to Community Preventive Services—methods. *American Journal of Preventive Medicine* 18(1, Supplement 1):35-43.

Brownson, R. C., J. G. Gurney, and G. H. Land. 1999. Evidence-based decision making in public health. *Journal of Public Health Management and Practice* 5(5):86-97.

Brownson, R. C., E. A. Baker, T. L. Leet, and K. N. Gillespie. 2003. *Evidence-based public health*. New York: Oxford University Press.

Brownson, R. C., C. Royer, R. Ewing, and T. D. McBride. 2006. Researchers and policymakers: Travelers in parallel universes. *American Journal of Preventive Medicine* 30(2):164-172.

Brownson, R. C., J. E. Fielding, and C. M. Maylahn. 2009a. Evidence-based public health: a fundamental concept for public health practice. *Annual Review of Public Health* 30(15):1-27.

Brownson, R. C., J. F. Chriqui, and K. A. Stamatakis. 2009b. Understanding evidence-based public health policy. *American Journal of Public Health* 99(9):1576-1583.

Cavill, N., C. Foster, P. Oja, and B. W. Martin. 2006. An evidence-based approach to physical activity promotion and policy development in Europe: Contrasting case studies. *Promotion & Education* 13(2):104-111.

Cheadle, A., T. D. Sterling, T. L. Schmid, and S. B. Fawcett. 2000. Promising community-level indicators for evaluating cardiovascular health-promotion programs. *Health Education Research* 15(1):109-116.

Choi, B. C. 2005. Twelve essentials of science-based policy. *Preventing Chronic Disease* 2(4):1-11.

Choi, B. C. K., T. Pang, V. Lin, P. Puska, G. Sherman, M. Goddard, M. J. Ackland, P. Sainsbury, S. Stachenko, H. Morrison, and C. Clottey. 2005. Can scientists and policy makers work together? *Journal of Epidemiology and Community Health* 59(8):632-637.

The Cochrane Collaboration. 2009. *The Cochrane Collaboration*. http://www.cochrane.org/ (accessed December 15, 2009).

Cook, D. J., D. L. Sackett, and W. O. Spitzer. 1995. Methodologic guidelines for systematic reviews of randomized control trials in health care from the Potsdam consultation on meta-analysis. *Journal of Clinical Epidemiology* 48(1):167-171.

Davidovitch, N., and D. Filc. 2006. Reconstructing data: Evidence-based medicine and evidence-based public health in context. *Dynamis* 26:287-306.

Dixon-Woods, M., R. Fitzpatrick, and K. Roberts. 2001. Including qualitative research in systematic reviews: Opportunities and problems. *Journal of Evaluation in Clinical Practice* 7(2):125-133.

Dobrow, M. J., V. Goel, and R. E. G. Upshur. 2004. Evidence-based health policy: Context and utilisation. *Social Science and Medicine* 58(1):207-217.

Dreisinger, M., T. L. Leet, E. A. Baker, K. N. Gillespie, B. Haas, and R. C. Brownson. 2008. Improving the public health workforce: Evaluation of a training course to enhance evidence-based decision making. *Journal of Public Health Management and Practice* 14(2):138-143.

Glasgow, R. E. 2008. What types of evidence are most needed to advance behavioral medicine? *Annals of Behavioral Medicine* 35(1):19-25.

Glasgow, R., L. Green, L. Klesges, D. Abrams, E. Fisher, M. Goldstein, L. Hayman, J. Ockene, and C. Olrleans. 2006. External validity: We need to do more. *Annals of Behavioral Medicine* 31(2):105-108.

Glasziou, P., and H. Longbottom. 1999. Evidence-based public health practice. *Australian and New Zealand Journal of Public Health* 23(4):436-440.

Green, L. W., and R. E. Glasgow. 2006. Evaluating the relevance, generalization, and applicability of research: Issues in external validation and translation methodology. *Evaluation & the Health Professions* 29(1):126-153.

Guyatt, G., and D. Rennie. 2002. *Users' guides to the medical literature: Essentials of evidence-based clinical practice*. Chicago, IL: American Medical Association Press.

Harris, M. R. 2000. Searching for evidence in perioperative nursing. *Seminars in Perioperative Nursing* 9(3):105-114.

Hausman, A. J. 2002. Implications of evidence-based practice for community health. *American Journal of Community Psychology* 30(3):453-467.

Hill, J. O., and J. C. Peters. 1998. Environmental contributions to the obesity epidemic. *Science* 280(5368):1371-1374.

Honore, P. A., R. L. Clarke, D. M. Mead, and S. M. Menditto. 2007. Creating financial transparency in public health: Examining best practices of system partners. *Journal of Public Health Management and Practice* 13(2):121-129.

IOM (Institute of Medicine). 2005. *Preventing childhood obesity: Health in the balance*. Edited by J. Koplan, C. T. Liverman and V. I. Kraak. Washington, DC: The National Academies Press.

IOM. 2007. *Progress in preventing childhood obesity: How do we measure up?* Edited by J. Koplan, C. T. Liverman, V. I. Kraak and S. L. Wisham. Washington, DC: The National Academies Press.

Jenicek, M. 1997. Epidemiology, evidenced-based medicine, and evidence-based public health. *Journal of Epidemiology* 7(4):187-197.

Jewell, C. J., and L. A. Bero. 2008. Developing good taste in evidence: Facilitators of and hindrances to evidence-informed health policymaking in state government. *Milbank Quarterly* 86(2):177-208.

Kerner, J. F. 2008. Integrating research, practice, and policy: What we see depends on where we stand. *Journal of Public Health Management and Practice* 14(2):193-198.

Khan, L. K., K. Sobush, D. Keener, K. Goodman, A. Lowry, J. Kakietek, and S. Zaro. 2009. Recommended community strategies and measurements to prevent obesity in the United States. *Morbidity and Mortality Weekly Report* 58(RR-7):1-26.

Klesges, L. M., D. A. Dzewaltowski, and R. E. Glasgow. 2008. Review of external validity reporting in childhood obesity prevention research. *American Journal of Preventive Medicine* 34(3):216-223.

Kohatsu, N. D., and R. J. Melton. 2000. A health department perspective on the Guide to Community Preventive Services. *American Journal of Preventive Medicine* 18(1, Supplement):3-4.

Kohatsu, N. D., J. G. Robinson, and J. C. Torner. 2004. Evidence-based public health: An evolving concept. *American Journal of Preventive Medicine* 27(5):417-421.

Kumanyika, S., and R. C. Brownson. 2007. *Handbook of obesity prevention: A resource for health professionals*. New York: Springer.

Kumanyika, S., R. W. Jeffery, A. Morabia, C. Ritenbaugh, and V. J. Antipatis. 2002. Obesity prevention: The case for action. *International Journal of Obesity* 26(3):425-436.

Laforest, R., and M. Orsini. 2005. Evidence-based engagement in the voluntary sector: Lessons from Canada. *Social Policy and Administration* 39(5):481-497.

Lytle, L. A. 2009. Measuring the food environment: State of the science. *American Journal of Preventive Medicine* 36(4, Supplement):134-144.

McGloin, A., L. Delaney, E. Hudson, and P. Wall. 2009. Symposium on "The challenge of translating nutrition research into public health nutrition." Session 5: Nutrition communication the challenge of effective food risk communication. *Proceedings of the Nutrition Society* 68(2):135-141.

McKinlay, J. B. 1992. Health promotion through healthy public policy: The contribution of complementary research methods. *Canadian Journal of Public Health* 83(Supplement 1):S11-S19.

McKinnon, R. A., J. Reedy, M. A. Morrissette, L. A. Lytle, and A. L. Yaroch. 2009. Measures of the food environment: A compilation of the literature, 1990-2007. *American Journal of Preventive Medicine* 36(4, Supplement):124-133.

Muir Gray, J. A. 2009. *Evidence-based healthcare and public health: How to make decisions about health services and public health*. 3rd ed. New York: Churchill Livingstone (Elsevier).

Pawson, R., T. Greenhalgh, G. Harvey, and K. Walshe. 2005. Realist review—a new method of systematic review designed for complex policy interventions. *Journal of Health Services Research and Policy* 10(Supplement 1):21-34.

Petticrew, M., and H. Roberts. 2003. Evidence, hierarchies, and typologies: Horses for courses. *Journal of Epidemiology and Community Health* 57(7):527-529.

Rychetnik, L., P. Hawe, E. Waters, A. Barratt, and M. Frommer. 2004. A glossary for evidence based public health. *Journal of Epidemiology and Community Health* 58(7):538-545.

Satterfield, J. M., B. Spring, R. C. Brownson, E. J. Mullen, R. P. Newhouse, B. B. Walker, and E. P. Whitlock. 2009. Toward a transdisciplinary model of evidence-based practice. *Milbank Quarterly* 87(2):368-390.

Swinburn, B., T. Gill, and S. Kumanyika. 2005. Obesity prevention: A proposed framework for translating evidence into action. *Obesity Reviews* 6(1):23-33.

West, S., J. Biesanz, and S. Pitts. 2000. Causal inference and generalization in field settings: Experimental and quasi-experimental designs. In *Handbook of research methods in social and personality psychology*, edited by H. Reis and C. Judd. Cambridge: University Press. Pp. 40-84.

4

Defining the Problem: The Importance of Taking a Systems Perspective

KEY MESSAGES

- Addressing the obesity epidemic requires complex thinking that to date has not been fully employed in prevention efforts.
- The demand for results-oriented research is gradually forcing a shift beyond linear models to models that also address the dynamic, multilevel complexity of real-world contexts.
- The systems approach is one such model. It encompasses the whole picture, highlights the broader context, considers interactions among multiple levels, recognizes the dynamic shifts that occur over time, and encourages collaboration among investigators from a variety of disciplines.

Obesity is a multifaceted problem that warrants complex thinking and a broad systems perspective to frame the problem, understand potential causes, identify critical leverage points of influence, and take effective action. Linear approaches to complex public health problems such as the obesity crisis are clearly useful, but cannot address the multiple dimensions of the real world and the many influences on the energy balance equation (Foresight, 2007). It is necessary to embrace complexity and to develop strategies and implement change at multiple levels to influence human behavior and reverse the current upward trends in weight. A systems perspective offers a new approach to obesity research and action that can meet this challenge (Huang et al., 2009).

The systems approach has a nearly 50-year history since its development by Forrester at the Massachusetts Institute of Technology (Forrester, 1991). Increasingly, obesity scholars are looking to other disciplines, from biology to psychology to computer sciences and engineering, that use this approach. In the health arena, the approach has been used to elucidate seemingly intractable problems, including cardiovascular disease (Homer et al., 2004), diabetes (Milstein et al., 2007), mental health (Smith et al., 2004), public health emergencies (Hoard et al., 2005), and tobacco control (National Cancer Institute, 2007).

The complex issue of obesity lends itself to a systems approach quite well. Like tobacco control, which employed diverse and multilevel strategies (Abrams et al., 2003, 2010), progress in the obesity field will require a paradigm shift toward an interdisciplinary knowledge base that integrates systems theory with concepts and practice from community development, social ecology, social networks, and public health (Best et al., 2003).

This chapter explains how systems thinking expands upon the multilevel, multisector strategies already proposed or in use to prevent obesity. It provides a primer on the concepts of such thinking and examines how the systems approach can be applied to identify the determinants, strategies, and actions that must be considered to address the obesity crisis. The chapter provides several practical examples of how systems thinking can be used in both small and large ways to expand the boundaries of current models and advance effective change in public health. The chapter also links the systems approach and its application to the L.E.A.D. framework (Figure 4-1), describing how it enhances the ability to generate, use, and learn from evidence and explaining how specific content pertaining to each step of the framework will differ according to the system on which one is focusing. Box 4-1 defines the key systems concepts pertinent to the discussion.

UNDERSTANDING A SYSTEMS APPROACH

As explained in Chapter 2, multilevel, multisector strategies, often based on ecological models (e.g., Figure 1-5 in Chapter 1),[1] have gained widespread acceptance for understanding the determinants of obesity and for framing prevention and control activities (Glass and McAtee, 2006). While these models acknowledge the multiple levels of a system and show their interrelationships, however, they may not always be complex enough to capture the dynamic interactions and the short- and long-term feedback loops among the many influences on the energy balance (Foresight, 2007; Sterman, 2006). Systems investigation can complement other methods by capturing this complexity, translating it into actions that can have an impact on the obesity problem and making it possible to predict unintended consequences and time-delayed effects (Mabry et al., 2008).

[1] An ecological model captures the many factors—individual, interpersonal, organizational, community (both social and economic), and health policy—that affect individual behavior change and therefore health.

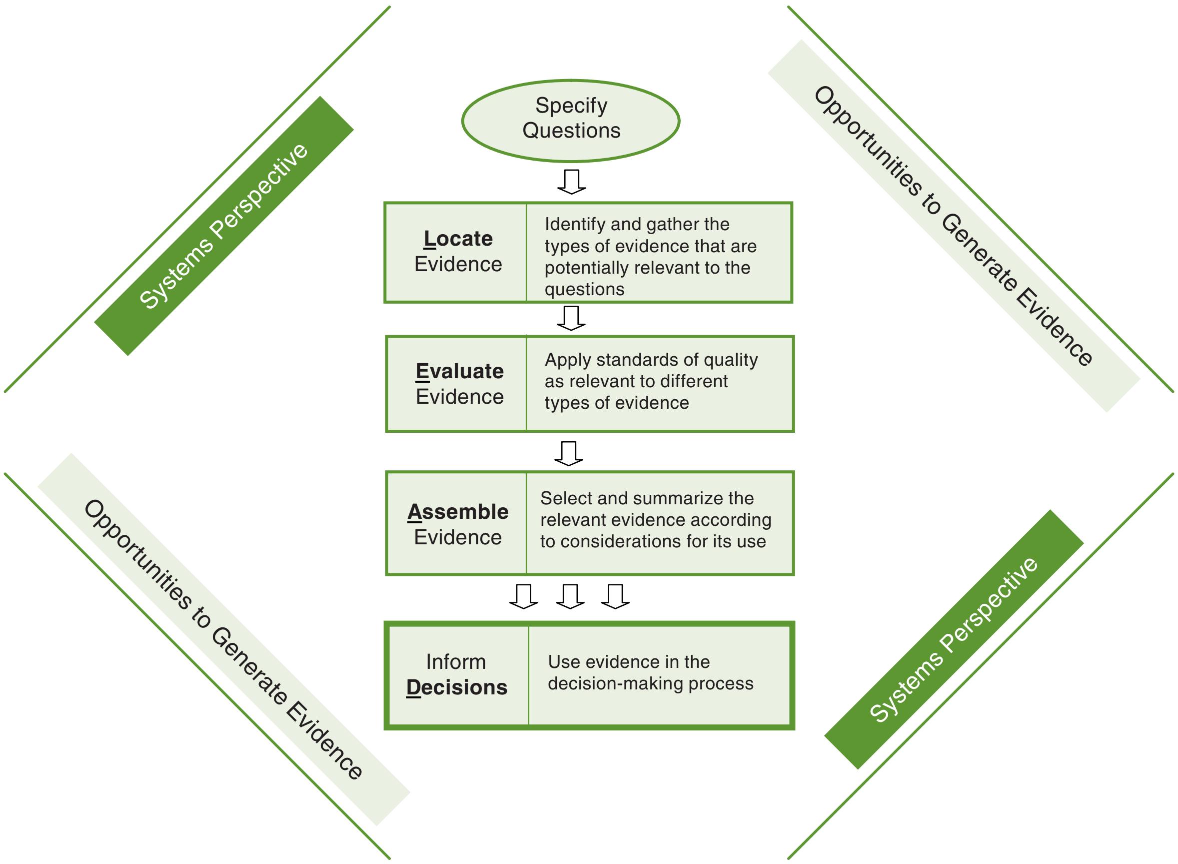


FIGURE 4-1 The **L**ocate Evidence, **E**valuate Evidence, **A**ssemble Evidence, Inform **D**ecisions (L.E.A.D.) framework for obesity prevention decision making.
NOTE: The element of the framework addressed in this chapter is highlighted.

A systems approach requires seeing the whole picture and not just a fragment, understanding the broader context, appreciating interactions among levels, and taking an interdisciplinary approach (Leischow and Milstein, 2006). A systems approach highlights the importance of the circumstances, or context, in which an action is taken in order to understand its implementation and potential impact. Thus while investigators must, for practical reasons, establish boundaries to define the system being studied, they must also recognize that each system exists within and interacts with a hierarchy of nested systems (Midgley, 2000). In addition, appreciating leverage points or points of power within a system can help explain how a small shift in one element of a complex system can produce larger changes in other elements (Meadows, 1999). These advantages of systems investigation are particularly important for interventions targeting obesity, given their far-reaching impact on the population; solutions should be designed to maximize benefit and minimize negative consequences.

The systems approach offers a further advantage with respect to the well-recognized gap between research and practice, which limits the extent to which advances in research translate to advances in improving public health. Most efforts to

Box 4-1
Definition of Key Systems Concepts

Systems approach: A paradigm or perspective involving a focus on the whole picture and not just a single element, awareness of the wider context, an appreciation for interactions among different components, and transdisciplinary thinking (Leischow and Milstein, 2006).

Systems investigation: A promising new frontier for research and action in response to complex and critical challenges (Leischow and Milstein, 2006).

Systems thinking: An iterative learning process in which one takes a broad, holistic, long-term, perspective of the world and examines the linkages and interactions among its elements (Sterman, 2006).

Systems theory: An interdisciplinary theory that requires merging of multiple perspectives and sources of information and deals with complex systems in technology, society, and science (Best et al., 2003).

Systems science: Research related to systems theory that offers insight into the nature of the whole system that often cannot be gained by studying the system's component parts in isolation (Mabry et al., 2008).

System dynamics: A methodology for mapping and modeling the forces of change in a complex system in order to better understand their interaction and govern the direction of the system; it enables stakeholders to combine input into a dynamic hypothesis that uses computer simulation to compare various scenarios for achieving change (Milstein and Homer, 2006).

link research to practice and policy have merely highlighted the challenges of transferring knowledge from single-discipline, highly controlled research to practice settings. Interdisciplinary investigation using a systems approach can potentially help close this gap (Mabry et al., 2008).

A systems approach to solving health problems requires new tools, including data, methods, theories, and statistical analysis different from those traditionally used in linear approaches. No single discipline can provide these tools. Therefore, it is necessary to approach health research with a collaborative team of investigators who bring knowledge and expertise from a variety of disciplines and sectors (Leischow et al., 2008). The theoretical frameworks and methodologies that result from such collaboration can generate new conceptual syntheses, new measurement techniques (e.g., social network analysis), and interdisciplinary fields of inquiry (e.g., behavioral genetics) with the capacity to tackle complex population health problems (Fowler et al., 2009).

Sterman (2006) explains how the dynamics of a system work, using policy resistance as an example. His explanation, reproduced in Box 4-2, encompasses the key concepts and variables in systems thinking: stocks, flows, feedback processes (positive or self-reinforcing and negative or self-correcting), side effects, and time delays.

USES OF SYSTEMS THINKING, APPROACHES, MAPPING, AND MODELING

Systems can be small or large and often coincide with the levels defined in an ecological model. For example, a school can be thought of as a micro-system within a larger community; as a meso-system within the even larger national, political, and social milieu; or as a macro-system within a global-system context. This section provides several examples of how systems thinking pertains to public health problems: body mass index (BMI) screening in schools, tobacco control in the United States, obesity modeling in the United States, and obesity prevention in the United Kingdom.

BMI Screening in Schools

The monitoring of childhood growth has been a contentious issue for several decades (James and Lobstein, 2009). In recent years, school districts have been under pressure to respond to the childhood obesity epidemic. Despite limited evidence on the value of schools providing individually directed help for children with higher BMIs, the establishment of school-based surveillance to document obesity prevalence and to inform the development of prevention and treatment policies has been recommended (e.g., Massachusetts Department of Public Health, 2009). Although the measurement of weight and height within schools appears relatively simple and in fact has been taking place for decades, the development of rigorous measurement and reporting protocols has been limited until recently. This lack of a well-defined process, together with the failure to take a systems perspective, can result in a number of unintended consequences and perturbations to the system. For example, children may feel embarrassed or stigmatized during the process, parents may feel unequipped to act on the information they receive, health care providers in the community may not be educated about obesity treatment, the community may lack adequate pediatric programming to which children can be referred, and schools may forego other screening programs to make room for BMI screening (for example, hearing and vision screenings were cut back when the Massachusetts Public Health Council voted to require BMI screening of schoolchildren [Mullen, 2009]). In addition, although obesity rates may be high, insufficient funds or a lack of political will may prevent the school system from accessing the funding and assistance needed to address these unintended consequences, leaving the community feeling frustrated and helpless.

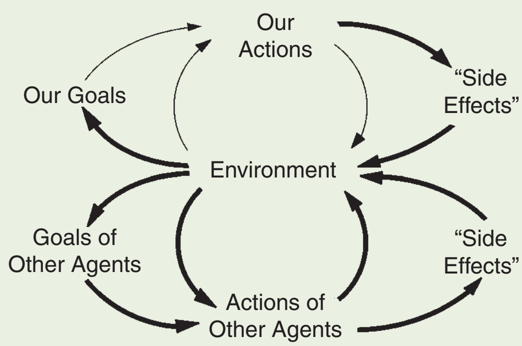
The Tobacco Control Movement

Although historically the tobacco control movement targeted individuals and their behaviors, it evolved into a multilevel systems approach to the problem (Abrams, 2007). Clearly, individual behavior change was the goal, but strategies involving industry, legislation, public health programming and messaging, and the health care system worked together to create that change. None of the strategies implemented as part of the movement worked alone. State by state, it was demonstrated that a combi-

Box 4-2
Primer on Concepts and Variables in Systems Thinking

Social systems contain intricate networks of feedback processes, both self-reinforcing (positive) and self-correcting (negative) loops. Failure to focus on feedback in policy design has critical consequences. Suppose the hospital you run faces a deficit, caught between rising costs and increasing numbers of uninsured patients. In response, you might initiate quality improvement programs to boost productivity, announce a round of layoffs, and accelerate plans to offer new high-margin elective surgical services. Your advisors and spreadsheets suggest that these decisions will cut costs and boost income. Problem solved—or so it seems.

Contrary to the open-loop model behind these decisions, the world reacts to our interventions (Figure). There is feedback: our actions alter the environment and, therefore, the decisions we take tomorrow. Our actions may trigger so-called side effects that we did not anticipate. Other agents, seeking to achieve *their* goals, act to restore the balance that we have upset; their actions also generate intended and unintended consequences. Goals are also endogenous, evolving in response to changing circumstances. For example, we strive to earn more in a quest for greater happiness, but habituation and social comparison rapidly erode any increase in subjective well-being (Kahneman et al., 1999).



Sources of policy resistance.

NOTE: Arrows indicate causation, e.g., our actions alter the environment. Thin arrows show the basic feedback loop through which we seek to bring the state of the system in line with our goals. Policy resistance (thick arrows) arises when we fail to account for the so called "side effects" of our actions, the responses of other agents in the system (and the unanticipated consequences of these), the ways in which experience shapes our goals, and the time delays often present in these feedbacks.

Policy resistance arises because we do not understand the full range of feedbacks surrounding—and created by—our decisions. The improvement initiatives you mandated never get off the ground because layoffs destroyed morale and increased the workload for the remaining employees. New services were rushed to market before all the kinks were worked out; unfavorable word of mouth causes the number of lucrative elective procedures to fall as patients flock to competitors. More chronically ill patients show up in your ER with complications after staff cuts slashed resources for patient education and follow-up; the additional workload forces still greater cuts in prevention. Stressed by long hours and continual crisis, your most experienced nurses and doctors leave for jobs with competitors, further raising the workload and undercutting quality of care. Hospital-acquired infections and preventable errors increase. Malpractice claims multiply. Yesterday's solutions become today's problems.

Ignoring the feedbacks in which we are embedded leads to policy resistance as we persistently react to the symptoms of difficulty, intervening at low leverage points and triggering delayed and distant effects. The problem intensifies, and we react by pulling those same policy levers still harder in an unrecognized vicious cycle. Policy resistance breeds cynicism about our ability to change the world for the better. Systems thinking requires us to see how our actions feed back to shape our environment. The greater challenge is to do so in a way that empowers, rather than reinforces, the belief that we are helpless victims of forces that we neither influence nor comprehend.

Time delays

Time delays in feedback processes are common and particularly troublesome. Most obviously, delays slow the accumulation of evidence. More problematic, the short- and long-run impacts of our policies are often different (smoking gives immediate pleasure, while lung cancer develops over decades). Delays also create instability and fluctuations that confound our ability to learn. Driving a car, drinking alcohol, and building a new semiconductor plant all involve time delays between the initiation of a control action (accelerating/braking, deciding to "have another," the decision to build) and its effects on the state of the system. As a result, decision makers often continue to intervene to correct apparent discrepancies between the desired and actual state of the system even after sufficient corrective actions have been taken to restore equilibrium. The result is overshoot and oscillation: stop-and-go traffic, drunkenness, and high-tech boom and bust cycles (Sterman, 1989). Public health systems are not immune to these dynamics, from oscillations in incidence of infectious diseases, such as measles (Anderson et al., 1984) and syphilis (Grassley et al., 2005), to the 2004-2005 flu vaccine fiasco, with scarcity and rationing followed within months by surplus stocks (U.S. House Committee on Government Reform, 2005).

Stocks and Flows

Stocks and the flows that alter them (the concepts of prevalence and incidence in epidemiology) are fundamental in disciplines from accounting to zoology: a population is increased by births and decreased by mortality; the burden of mercury in a child's body is increased by ingestion and decreased by excretion. The movement and transformation of material among states is central to the dynamics of complex systems. In physical and biological systems, resources are usually tangible: the stock of glucose in the blood; the number of active smokers in a population. The performance of public health systems, however, is also determined by resources such as physician skills, patient knowledge, community norms, and other forms of human, social, and political capital.

Research shows people's intuitive understanding of stocks and flows is poor in two ways. First, narrow mental model boundaries mean that people are often unaware of the networks of stocks and flows that supply resources and absorb wastes. California's Air Resources Board seeks to reduce air pollution by promoting so-called zero emission vehicles (California Air Resources Board, 2010). True, zero emission vehicles need no tailpipe. But the plants required to make the electricity or hydrogen to run them do generate pollution. California is actually promoting *displaced* emission vehicles, whose wastes would blow downwind to other states or accumulate in nuclear waste dumps outside its borders. Air pollution causes substantial mortality, and fuel cells may prove to be an environmental boon compared with internal combustion. But no technology is free of environmental impact, and no legislature can repeal the second law of thermodynamics.

Second, people have poor intuitive understanding of the process of accumulation. Most people assume that system inputs and outputs are correlated (e.g., the higher the federal budget deficit, the greater the national debt will be) (Booth Sweeny and Sterman, 2000). However, stocks integrate (accumulate) their net inflows. A stock rises even as its net inflow falls, as long as the net inflow is positive: the national debt rises even as the deficit falls—debt falls only when the government runs a surplus; the number of people living with HIV continues to rise even as incidence falls—prevalence falls only when infection falls below mortality. Poor understanding of accumulation has significant consequences for public health and economic welfare. Surveys show most Americans believe climate change poses serious risks, but they also believe that reductions in GHG emissions sufficient to stabilize atmospheric GHG concentrations can be deferred until there is greater evidence that climate change is harmful (Sterman and Booth Sweeny, 2007). Federal policy makers likewise argue that it is prudent to wait and see whether climate change will cause substantial economic harm before undertaking policies to reduce emissions (Bush, 2002). Such wait-and-see policies erroneously presume that climate change can be reversed quickly should harm become evident, underestimating immense delays in the climate's response to GHG emissions. Emissions are now about twice the rate at which natural processes remove GHGs from the atmosphere (Houghton et al., 2001). GHG concentrations will therefore continue to rise even if emissions fall, stabilizing only when emissions equal removal. In contrast, experiments with highly educated adults—graduate students at MIT—show that most believe atmospheric GHG concentrations can be stabilized while emissions into the atmosphere continuously exceed the removal of GHGs from it (Sterman and Booth Sweeny, 2007). Such beliefs are analogous to arguing that a bathtub filled faster than it drains will never overflow. They violate conservation of matter, and the violation matters: wait-and-see policies guarantee that atmospheric GHG concentrations, already greater than any in the past 420,000 years (Houghton et al., 2001), will rise far higher, increasing the risk of dangerous changes in climate that may significantly harm public health and human welfare.

NOTE: ER = emergency room; GHG = greenhouse gas; MIT = Massachusetts Institute of Technology.
SOURCE: Excerpt from Sterman, 2006. Reprinted with permission of the American Public Health Association. Sterman, J.D., Learning from evidence in a complex world, *American Journal of Public Health* 96(3) 505-513, 2006, copyright © 2006 by the American Public Health Association.

nation of strategies was better than any single intervention, and the more components that were used, the better (CDC, 1999; Levy et al., 2007). The movement can be seen as a good example to inform the field of obesity prevention.

To explore this paradigm shift, a pilot project, the Initiative on the Study and Implementation of Systems (ISIS), was conducted with funding from the National Cancer Institute (NCI) (Best et al., 2006). ISIS was designed to (1) explore how systems thinking approaches might improve understanding of the factors contributing to tobacco use, (2) inform strategic decision making on which efforts might be most effective for reducing tobacco use, and (3) serve as an exemplar for addressing other public health problems. Contextually, tobacco control can be viewed as a system comprising smaller systems and existing within the broader systems of public health; economics; and society at the local, regional, and global levels. Figure 4-2 shows the evolution of tobacco control approaches toward systems thinking (NCI, 2007).

The ISIS project was an attempt to understand the whole problem of tobacco use comprehensively, and ultimately to address the problem through systems change. As a result of strategic planning, the ISIS group identified four priority areas (Figure 4-3) that together serve as a synergistic foundation for understanding and improving public health from a systems perspective (Leischow et al., 2008; NCI,

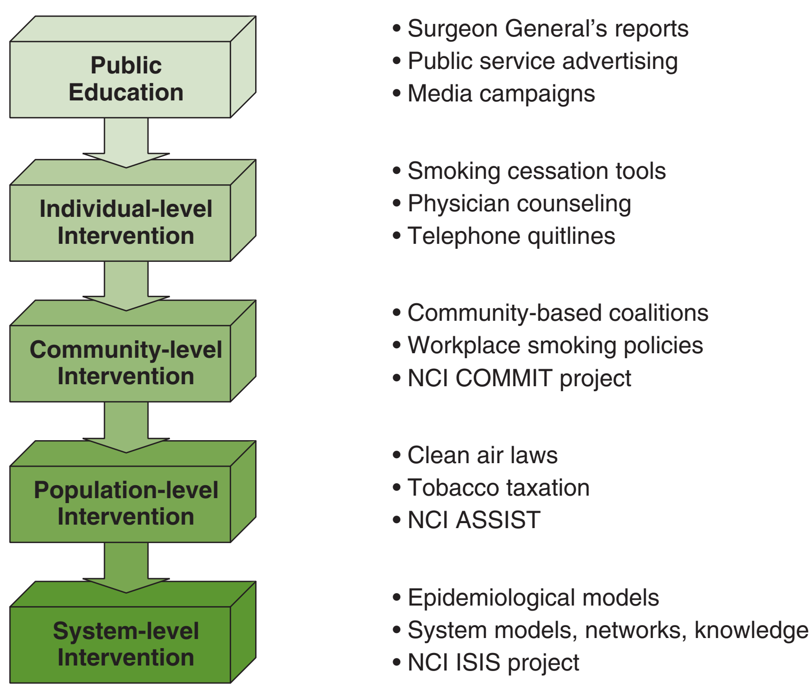


FIGURE 4-2 Evolution of tobacco control approaches toward systems thinking.
NOTES: Quitlines indicate telephone hotlines for smoking cessation. ASSIST = American Stop Smoking Intervention Study for Cancer Prevention; COMMIT = Community Intervention Trial for Smoking Cessation; ISIS = Initiative on the Study and Implementation of Systems; NCI = National Cancer Institute.
SOURCE: NCI, 2007.

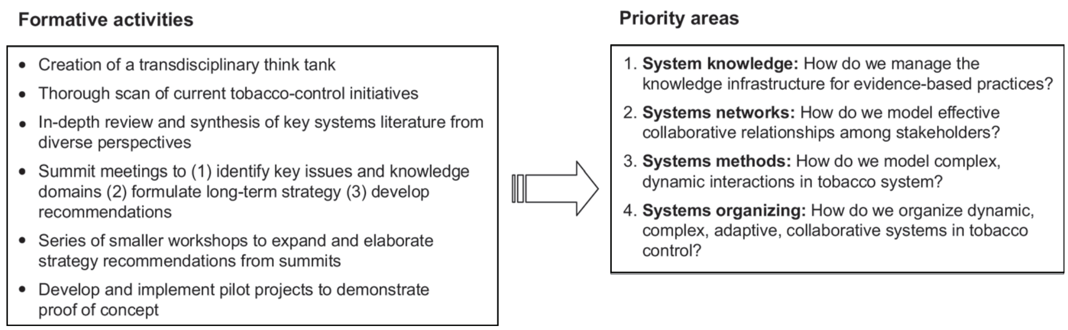


FIGURE 4-3 Priority areas identified by the Initiative on the Study and Implementation of Systems (ISIS) group.
SOURCE: Reprinted from Leischow et al., Copyright © 2008, with permission from Elsevier.

2007). Simulation modeling, conducted in the third of ISIS's priority areas, has been useful for exploring the impact of changes in various parameters within the complex systems that have an impact on population-level outcomes related to tobacco use behaviors and policies (Abrams et al., 2010; Levy et al., 2010a,b).

Mapping of Obesity Causality in the United States

System dynamics modeling can help map causality by addressing risks and outcomes; when performed prospectively, it can be used to predict future outcomes, and when performed retrospectively, it can be used to understand how strategies and delivery systems interacted with a population during an intervention. The process proceeds iteratively through the general steps shown in Figure 4-4, beginning with the identification of a persistent problem (Milstein and Homer, 2006). Milstein and Homer

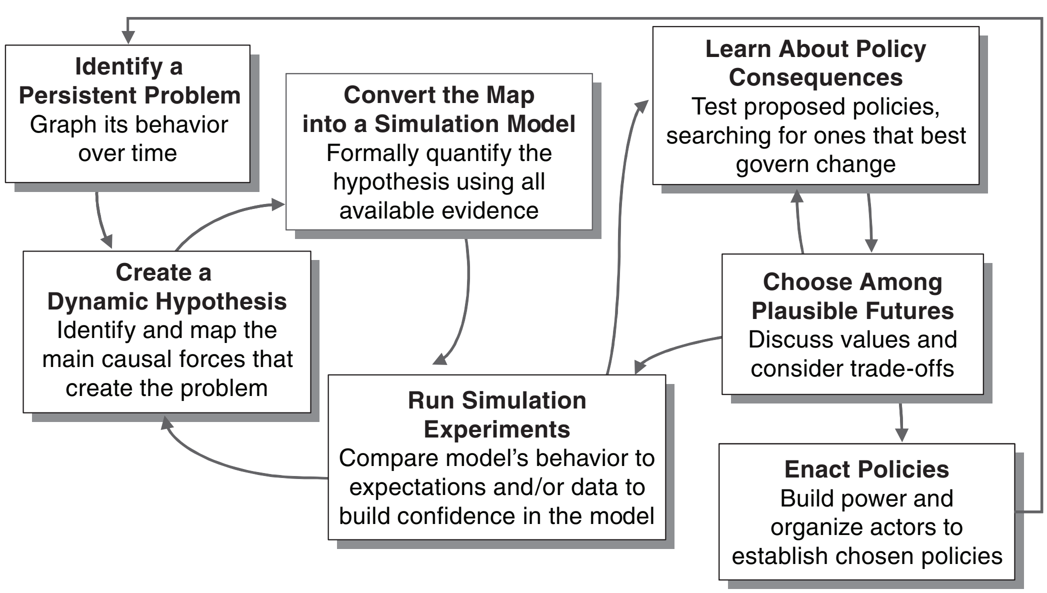


FIGURE 4-4 Iterative steps in system dynamics modeling.
SOURCE: Milstein and Homer, 2006.

(2009) use system dynamics modeling to map the forces that contribute to the persistent obesity problem (see Figure 4-5). This exercise helps in understanding the causes of obesity, the broader systems to which they belong, and how they are thought to interact. Such mapping can therefore elucidate potential mechanisms and dynamic pathways on which intervention strategies should focus.

Once such maps have been developed, they can be converted into computer simulation models that can be used to identify interventions and policies with the potential to alleviate the problem. These experiments are followed by sensitivity analyses to assess areas of uncertainty in the models and to guide future research. Once the models have been finalized, stakeholders are convened to participate in "action labs," in which the models allow them to discover for themselves the likely consequences of alternative policy scenarios (Milstein and Homer, 2009).

A broad array of modeling techniques are used for different purposes in many different fields. Because of the complexity of the obesity problem, the most suitable modeling techniques will have several characteristics (Hammond, 2009). First, because of the scale of the epidemic, models may provide the most insight if they capture multiple levels of analysis. Second, to capture the dynamics of such a complex system, models must be able to incorporate individual heterogeneity and adaptation over time. Finally, models must be able not only to provide a better understanding of the problem and the mechanisms behind it, but also to aid in the design of new and better interventions to slow and reverse the epidemic.

Obesity Prevention in the United Kingdom

The most comprehensive effort to both understand and map the obesity epidemic and formulate a national action plan was carried out by the Foresight Group in the United Kingdom (Foresight, 2007). The Foresight Tackling Obesities: Future Choices Project was aimed at producing a sustainable response to obesity in the United Kingdom over the 40 years following the plan's release. Using a systems approach, the group pursued objectives that included using the scientific evidence base across a wide range of disciplines to identify the many factors that influence obesity, looking beyond the obvious to achieve an integrated understanding of the relationships among these factors and their relative importance, building on this evidence to identify effective interventions, analyzing how future levels of obesity might change, and identifying what the most effective future responses might be. A detailed causal loop obesity system map was produced to display the interrelationships among the various contributors to energy balance; a simplified version of this map is shown in Figure 4-6.

In addition, a strategic framework was developed to identify gaps in current initiatives that would have to be filled to mount an integrated policy response. Its authors identify six key elements of this framework (Jebb, 2009):

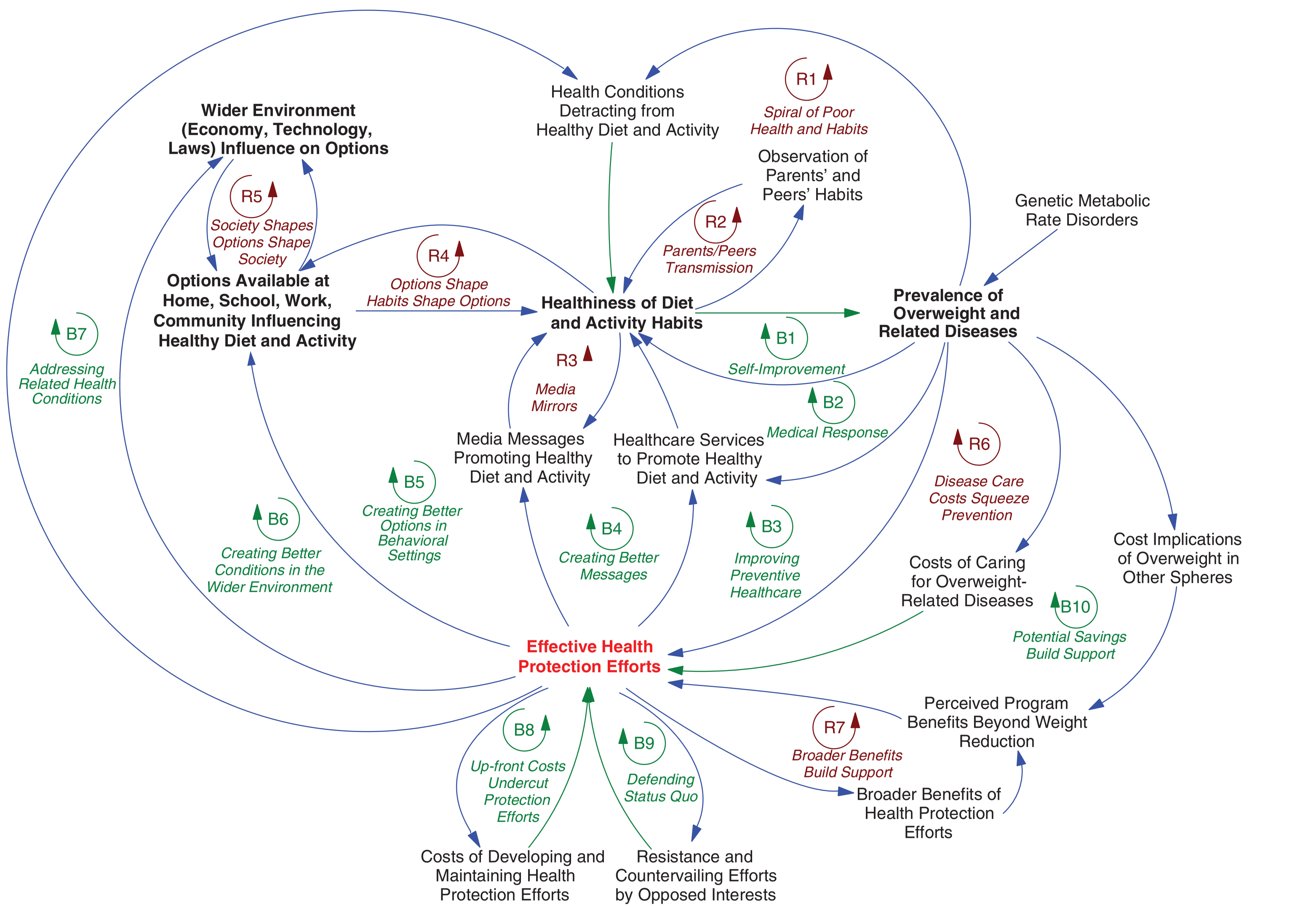


FIGURE 4-5 The obesity "system": a broad causal map.

NOTES: Blue arrows indicate same-direction links; green arrows indicate opposite-direction links; R loops indicate reinforcing processes; B loops indicate balancing processes. All parameters vary by such factors as age, sex, race/ethnicity, income, and geography.

SOURCE: Milstein and Homer, 2009. Reprinted with permission.

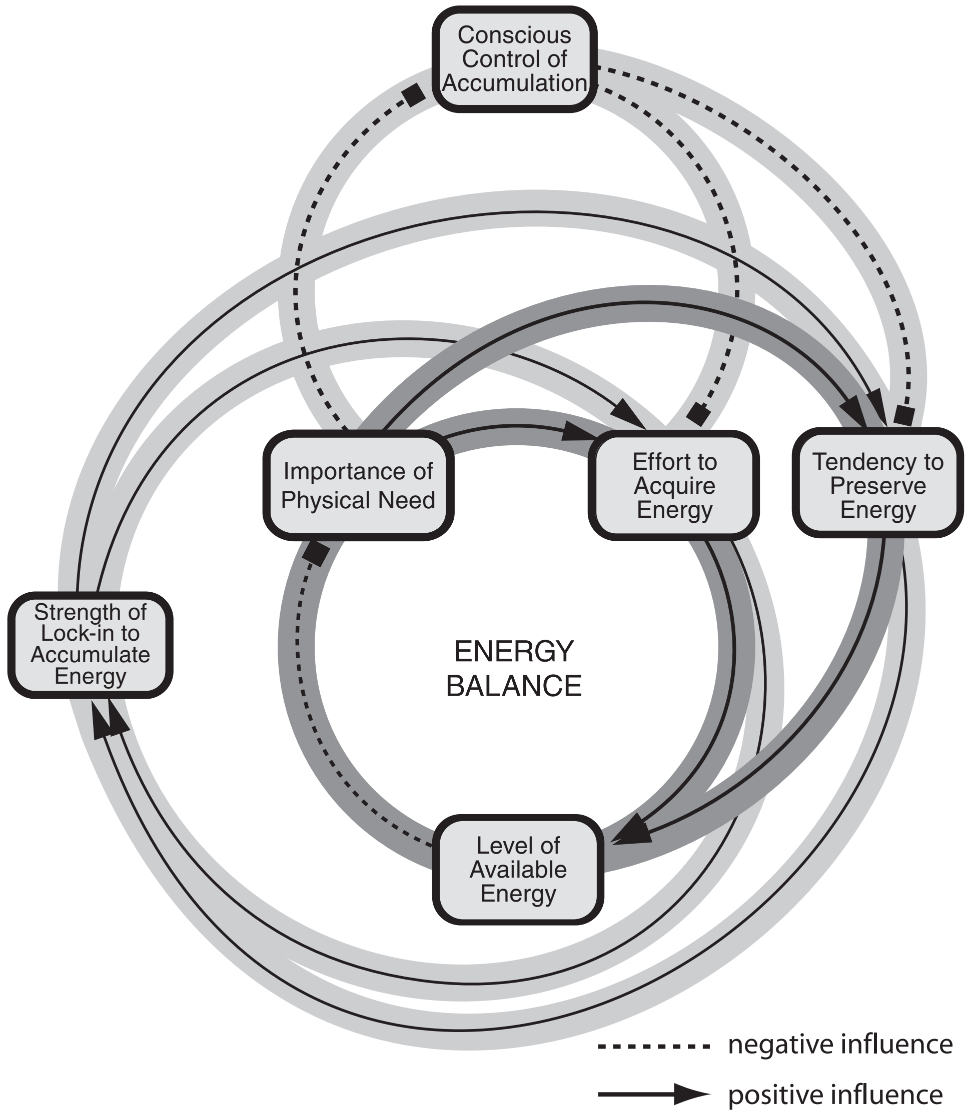


FIGURE 4-6 Simplified version of the causal loop obesity system map showing the interrelationships among various contributors to energy balance. The map was developed by the UK Foresight Group to understand and chart the obesity epidemic in order to inform a national action plan.
SOURCES: Vandenbroeck et al., 2007.

- systematic change that addresses the diverse determinants of obesity simultaneously to minimize the risk of compensatory actions;
- integrated interventions at all levels of society—individual, family, local, national and international—recognizing that individual choices are shaped by the wider context;
- interventions across the life course to reinforce and sustain long-term behavior change;

- diverse interventions that combine focused initiatives (which impose change), "enablers" (which inform and facilitate change), and "amplifiers" (which address social norms and the cultural context);
- actions planned over time such that early initiatives build a climate for subsequent interventions; and
- ongoing evidence gathering, including population-level surveillance and evaluation of interventions.

The Foresight Group identified as a key research challenge "the evaluation of new policy initiatives at all levels (process audits, behaviors and biomarkers, long-term health and economic outcomes)" (Jebb, 2009, p. 39). These types of causal loops magnify the need for new ways of evaluating and incorporating evidence not only from research studies, but also from the real-world experience of obesity initiatives undertaken not just within a particular country but around the world.

Although outcomes of systems approaches such as the Foresight Group's causal loop system map appear complex, they are useful for informing practical, real-world intervention strategies. For example, Foresight's Tackling Obesities: Future Choices Project was used to inform a cross-government strategy for England that was part of a sustained program to reduce obesity and support healthy weight maintenance (Cross-Government Obesity Unit, 2010).

RELATION TO THE L.E.A.D. FRAMEWORK

The idea that evidence should be identified, evaluated, and summarized from a systems perspective is fundamental to the framework proposed in this report. A systems perspective broadens the traditional approach to locating, evaluating, and assembling evidence (which generally limits the evidence to results of rigorous randomized controlled trials) to encompass evidence that reflects the complexity of the problem. Users of the framework are encouraged to approach every aspect of decision making with a comprehensive lens, considering the complex context in which programs and policies will be implemented and how it may affect their implementation and impact. A systems perspective enables the decision maker to understand interactions among smaller systems within the larger system and identify potential synergies or harms that should be explored before implementation. Creating "what if" scenarios based on systems maps can help decision makers and stakeholders think about various approaches and where to focus efforts, as well as potential costs; elements that are critical in the design of interventions or program and policy evaluations; and feedback loops that can be sources of evaluation data (see Chapter 6).

APPLICATION AND FUTURE DIRECTIONS

According to Hammond (2009), the most effective models for addressing the obesity epidemic are likely to be those that capture multiple mechanisms at multiple levels, integrate micro and macro data and dynamics, account for significant heterogeneities, and allow for policy experimentation. To fully realize the potential application of systems theory to obesity prevention, a number of strategies will be required. First, current and future leaders should be trained in the science and understanding of systems and their application to the obesity crisis. This training would include causal mapping, conceptualization of interventions, and computational and simulation modeling techniques. The application of these methods to the obesity epidemic will be challenging—there will be important data that are not yet available, uncertainty about a number of assumptions, and many key mechanisms whose inner workings are unknown (Hammond, 2009). In some cases, smaller systems will have to be studied independently, perhaps with relatively homogeneous populations, before being integrated into a more comprehensive model. Various combinations of models can then be explored and tested against the same outcome data, building slowly toward a model that encompasses the full breadth of the system by integrating all those narrower models. Second, empirical research should be funded and executed using systems theory as a guide. Focused studies can be used to confirm and quantify relationships and to test their effects. Ideally, this research would be carried out in conjunction with modeling studies to produce the most informative data and to guide future research. Third, both knowledge generators and users must work collaboratively with different disciplines to build interdisciplinary capacity.

Caution will be necessary regarding the use of models and the need to link their application with empirical research. The interplay between systems theory and research requires high-quality experimental and quasi-experimental designs. Systems thinking puts researchers in a better position to ask the right questions. Research applications allow a systems model to make the right predictions.

In conclusion, systems thinking in public health cannot be encompassed by a single discipline or even a single "systems thinking" approach (e.g., system dynamics models). Rather, it requires interdisciplinary integration of approaches to public health aimed at understanding and reconciling linear and nonlinear, qualitative and quantitative, and reductionistic and holistic thinking and methods to create a federation of systems approaches (NCI, 2007).

REFERENCES

Abrams, D. B. 2007. Comprehensive smoking cessation policy for all smokers: systems integration to save lives and money. In *Ending the tobacco problem: A blueprint for the nation*, edited by R. J. Bonnie, K. Stratton, and R. B. Wallace. Washington, DC: The National Academies Press. Pp. A1-A50.

Abrams, D. B., F. Leslie, R. Mermelstein, K. Kobus, and R. R. Clayton. 2003. Transdisciplinary tobacco use research. *Nicotine and Tobacco Research* 5(Supplement 1):S5-S10.

Abrams, D. B., A. L. Graham, D. T. Levy, P. L. Mabry, and C. T. Orleans. 2010. Boosting population quits through evidence-based cessation treatment and policy. *American Journal of Preventive Medicine* 38(3, Supplement 1):S351-S363.

Anderson, R. M., B. T. Grenfell, and R. M. May. 1984. Oscillatory fluctuations in the incidence of infectious disease and the impact of vaccination: Time series analysis. *Journal of Hygiene (London)* 93(3):587-608.

Best, A., G. Moor, B. Holmes, P. I. Clark, T. Bruce, S. Leischow, K. Buchholz, and J. Krajnak. 2003. Health promotion dissemination and systems thinking: Towards an integrative model. *American Journal of Health Behavior* 27(Supplement 3).

Best, A., R. Tenkasi, W. Trochim, F. Lau, B. Holmes, T. Huerta, S. Moor, S. Leischow, and P. Clark. 2006. Systemic transformational change in tobacco control: An overview of the Initiative for the Study and Implementation of Systems (ISIS). In *Innovations in health care: A reality check*, edited by A. L. Casebeer, A. Harrison, and A. L. Mark. New York: Palgrave Macmillan.

Booth Sweeney, L., and J. D. Sterman. 2000. Bathtub dynamics: Initial results of a systems thinking inventory. *System Dynamics Review* 16:249-294.

Bush, G. W. 2002. *President announces clear skies and global climate change initiatives.* http://georgewbush-whitehouse.archives.gov/news/releases/2002/02/20020214-5.html (accessed March 16, 2010).

California Air Resources Board. 2010. *Zero emission vehicle program.* http://www.arb.ca.gov/msprog/zevprog/zevprog.htm (accessed March 16, 2010).

CDC (Centers for Disease Control and Prevention). 1999. *Best practices for comprehensive tobacco control programs—April 1999.* Atlanta, GA: U.S. Department of Health and Human Services, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health.

Cross-Government Obesity Unit. 2010. *Healthy weight, healthy lives: Two years on.* http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/documents/digitalasset/dh_113495.pdf (accessed March 4, 2010).

Foresight. 2007. *Tackling obesities: Future choice—Project report, 2nd edition.* London: U.K. Government Office for Science.

Forrester, J. W. 1991. System dynamics and the lessons of 35 years. In *The systemic basis of policy making in the 1990s*, edited by K. B. De Greene. Cambridge, MA: MIT Press.

Fowler, J. H., C. T. Dawes, and N. A. Christakis. 2009. Model of genetic variation in human social networks. *Proceedings of the National Academy of Sciences of the United States of America* 106(6):1720-1724.

Glass, T. A., and M. J. McAtee. 2006. Behavioral science at the crossroads in public health: Extending horizons, envisioning the future. *Social Science and Medicine* 62(7):1650-1671.

Grassley, N. C., C. Fraser, and G. P. Garnett. 2005. Host immunity and synchronized epidemics of syphilis across the United States. *Nature* 433(7024):417-421.

Hammond, R. A. 2009. Complex systems modeling for obesity research. *Preventing Chronic Disease* 6(3):A96.

Hoard, M., J. Homer, W. Manley, P. Furbee, A. Haque, and J. Helmkamp. 2005. Systems modeling in support of evidence-based disaster planning for rural areas. *International Journal of Hygiene and Environmental Health* 208(1-2):117-125.

Homer, J., G. Hirsch, M. Minniti, and M. Pierson. 2004. Models for collaboration: How system dynamics helped a community organize cost-effective care for chronic illness. *System Dynamics Review* 20(3):199-222.

Houghton, J., Y. Ding, D. Griggs, M. Noguer, P. J. van der Linden, X. Dai, K. Maskell, and C. A. Johnson. 2001. *Climate change 2001: The scientific basis.* Cambridge, UK: Cambridge University Press.

Huang, T. T., A. Drewnowski, S. K. Kumanyika, and T. A. Glass. 2009. A systems-oriented multilevel framework for addressing obesity in the 21st century. *Preventing Chronic Disease* 6(3):A82.

James, W. P. T., and T. Lobstein. 2009. BMI screening and surveillance: An international perspective. *Pediatrics* 124(Supplement 1).

Jebb, S. 2009. Developing a strategic framework to prevent obesity. Presented at the IOM Workshop on the Application of Systems Thinking to the Development and Use of Evidence in Obesity Prevention Decision-Making, March 16, Irvine, CA.

Kahneman, D., E. Diener, and N. Schwarz. 1999. *Well-being: The foundations of hedonic psychology.* New York: Russel Sage.

Leischow, S. J., and B. Milstein. 2006. Systems thinking and modeling for public health practice. *American Journal of Public Health* 96(3):403-405.

Leischow, S. J., A. Best, W. M. Trochim, P. I. Clark, R. S. Gallagher, S. E. Marcus, and E. Matthews. 2008. Systems thinking to improve the public's health. *American Journal of Preventive Medicine* 35(2, Supplement 1):S196-S203.

Levy, D. T., H. Ross, L. Powell, J. E. Bauer, and H. R. Lee. 2007. The role of public policies in reducing smoking prevalence and deaths caused by smoking in Arizona: Results from the Arizona tobacco policy simulation model. *Journal of Public Health Management and Practice* 13(1):59-67.

Levy, D. T., A. L. Graham, P. L. Mabry, D. B. Abrams, and C. T. Orleans. 2010a. Modeling the impact of smoking-cessation treatment policies on quit rates. *American Journal of Preventive Medicine* 38(3, Supplement 1).

Levy, D. T., P. L. Mabry, A. L. Graham, C. T. Orleans, and D. B. Abrams. 2010b. Reaching Healthy People 2010 by 2013: A SimSmoke simulation. *American Journal of Preventive Medicine* 38(3, Supplement 1).

Mabry, P. L., D. H. Olster, G. D. Morgan, and D. B. Abrams. 2008. Interdisciplinarity and systems science to improve population health. A view from the NIH Office of Behavioral and Social Sciences Research. *American Journal of Preventive Medicine* 35(Supplement 2).

Massachusetts Department of Public Health. 2009. *BMI screening guidelines for schools.* http://www.mass.gov/Eeohhs2/docs/dph/mass_in_motion/community_school_screening.pdf (accessed April 5, 2010).

Meadows, D. 1999. *Leverage points: Places to intervene in a system.* Hartland, VT: The Sustainability Institute.

Midgley, G. 2000. *Systemic intervention: Philosophy, methodology and practice.* New York: Kluwer Academic/Plenum.

Milstein, B., and J. Homer. 2006. *Background on system dynamics simulation modeling with a summary of major public health studies.* Atlanta, GA: CDC. http://www.caldiabetes.org/get_file.cfm?contentID=501&ContentFilesID=389 (accessed February 3, 2010).

Milstein, B., and J. Homer. 2009. System dynamics simulation in support of obesity prevention decision-making. Presented at the Institute of Medicine Workshop on the Application of Systems Thinking to the Development and Use of Evidence in Obesity Prevention Decision-Making, March 16, Irvine, CA. http://www.iom.edu/~/media/Files/Activity%20Files/

PublicHealth/ObesFramework/IOMIrvine16Mar09v52MilsteinHomer.ashx (accessed April 20, 2010).

Milstein, B., A. Jones, J. B. Homer, D. Murphy, J. Essien, and D. Seville. 2007. Charting plausible futures for diabetes prevalence in the United States: A role for system dynamics simulation modeling. *Preventing Chronic Disease* 4(3):1-8.

Mullen, J. 2009. *Proposed Amendments to 105 CMR 200.000: Physical Examination of School Children*. Letter to Massachusetts Public Health Council from Director, Bureau of Community Health Access and Promotion, Massachusetts Department of Public Health.

NCI (National Cancer Institute). 2007. *Greater than the sum: Systems thinking in tobacco control. Tobaco control monograph no. 18*. NCI Tobacco Control Monograph Series. NIH Pub. No. 06-6085. Bethesda, MD: U.S. Department of Health and Human Services, National Institutes of Health, National Cancer Institute.

Smith, G., E. F. Wolstenholme, D. McKelvie, and D. Monk. 2004. *Using system dynamics in modeling mental health issues in the UK*. Paper presented at 22nd International Conference of the System Dynamics Society, Oxford, England.

Sterman, J. D. 1989. Modeling managerial behavior: Misperceptions of feedback in a dynamic decision making experiment. *Management Science* 35:321-339.

Sterman, J. D. 2006. Learning from evidence in a complex world. *American Journal of Public Health* 96(3):505-514.

Sterman, J. D., and L. Booth Sweeney. 2007. Understanding public complacency about climate change: Adults' mental models of climate change violate conservation of matter. *Climatic Change* C80(3-4):213-238.

U.S. House Committee on Government Reform. 2005. *The perplexing shift from shortage to surplus: Managing this season's flu shot supply and preparing for the future*. 1st Session, 109th Congress, February 10, 2005.

Vandenbroeck, I. P., J. Goossens, and M. Clemens. 2007. *Obesity system atlas*. Foresight tackling obesities: Future choices.

5

Specifying Questions and Locating Evidence: An Expanded View

KEY MESSAGES

- The L.E.A.D. framework advocates for an expanded perspective in identifying the forms of evidence that may be potentially relevant to obesity prevention decisions. Evidence to inform decisions about obesity prevention may come from a variety of disciplines, including other areas of public health. From this expanded perspective, there may be multiple potential sources of and ways to gather evidence.
- The framework includes an evidence typology that helps link questions related to obesity prevention decisions to different methods for locating evidence.
- To obtain a broad-based pool of evidence, questions may be specified at any time during the decision-making process and fed back to particular steps in the L.E.A.D. framework.
- Diverse methods used by different disciplines can be applied to answer a range of obesity prevention questions.
- Specialized research training may be required for the use of certain types of evidence.
- The process of locating evidence will reveal evidence gaps and highlight areas in which new evidence needs to be generated. Documenting these evidence gaps is critical for stimulating evaluation studies and other types of research focused on previously unexamined obesity prevention questions.

A basic premise of the L.E.A.D. framework is that good decisions should draw on a sound body of relevant evidence, acquired by applying a systematic process for locating, appraising, and compiling information relevant to a major public health issue such as obesity. Evidence-based policy and practice imply that interventions aimed at reducing obesity or its risk in the general population are supported by the best available research evidence on their efficacy and population impact. Decision makers also frequently seek additional information that has direct utility within their local context. Such information may be related to costs, implementation, scalability, sustainability, and factors that could facilitate or impede the success of an obesity prevention intervention. Locating useful evidence requires a clear concept of the types of information that may be useful for a particular purpose, as well as an awareness of where that information can be found.

This chapter addresses fundamental issues related to the types of evidence useful for answering a range of user questions and how that evidence can be located and gathered (Figure 5-1). Criteria for screening evidence are also addressed as background for the discussion of evidence evaluation in Chapter 6. The chapter begins by presenting an evidence typology to demonstrate how particular questions can be addressed with various methods. It then reviews potentially useful sources of evidence. The chapter concludes with a discussion of considerations in gathering the evidence.

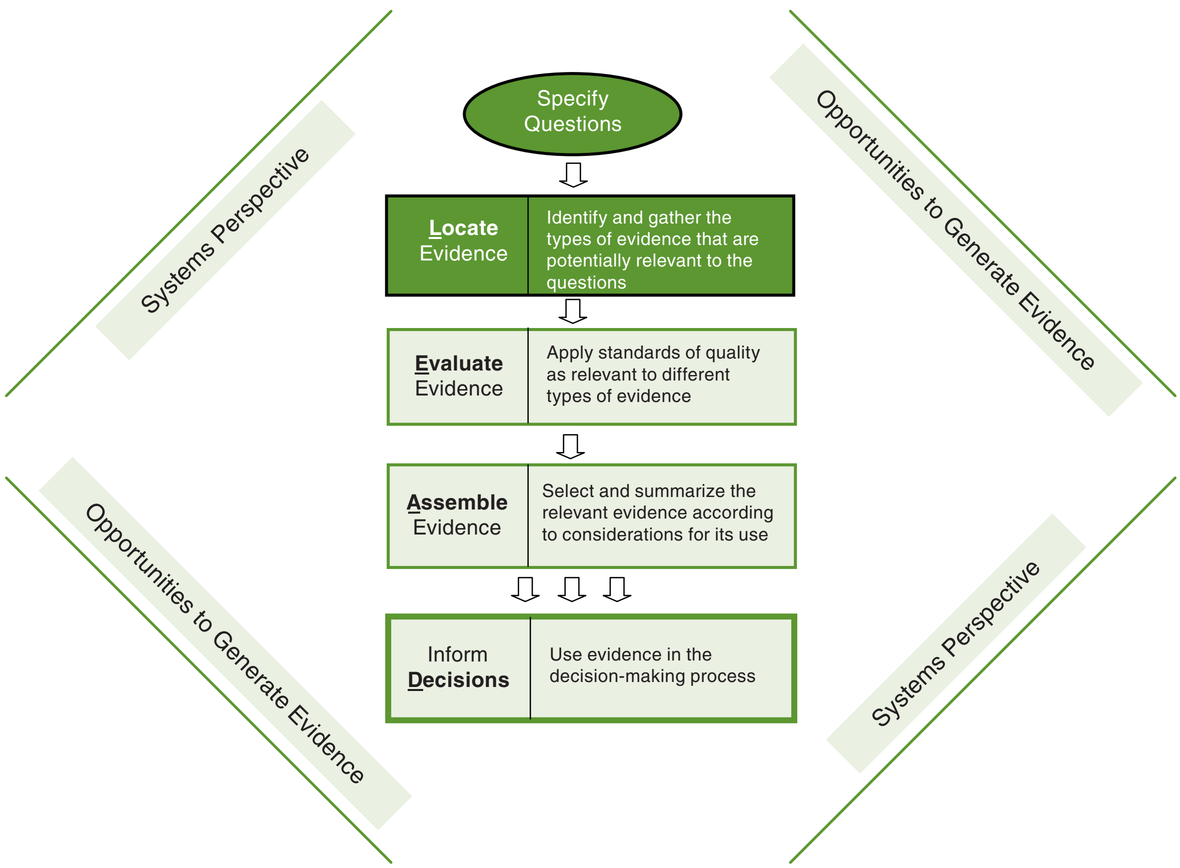


FIGURE 5-1 The **L**ocate Evidence, **E**valuate Evidence, **A**ssemble Evidence, Inform **D**ecisions (L.E.A.D.) framework for obesity prevention decision making.
NOTE: The elements of the framework addressed in this chapter are highlighted.

SPECIFYING QUESTIONS: AN EVIDENCE TYPOLOGY FOR THE L.E.A.D. FRAMEWORK

As explained in Chapter 3, the L.E.A.D. framework is oriented to decision makers and researchers who are typically called upon to use evidence to make decisions that apply within and across different levels of a system. The system may be at the local, regional, or national level; it may be a health or education system, another type of institution, a whole community, or a subsystem of a community. Evidence needs may vary substantively depending on when, where, at what level, and under what conditions users of the evidence operate. Differences in decision contexts are even found frequently within the same system. Evidence should be appropriately matched to user needs.

Locating a comprehensive pool of evidence useful for decision making is considerably facilitated when one starts with a focused and well-formulated set of guiding questions. The overall premise is that a problem is best grasped when examined from multiple perspectives with multiple forms of evidence, appropriately aligned with specific questions. Accordingly, the evidence typology of the L.E.A.D. framework identifies a set of guiding questions, adapting the approaches of the International Obesity Task Force framework and the evidence-based public health concepts outlined in Chapter 3 (see Figure 5-2): "Why" questions as shorthand for "Why should we do something about this problem in our situation?"; "What" questions as shorthand for "What specifically should we do about this problem?"; and "How" questions as shorthand for "How do we implement this information for our situation?" This typology is detailed in the following subsections.

"Why" Questions

Locating evidence guided by "Why" questions helps decision makers characterize the reasons for taking action on a public health issue in their particular region or locale.

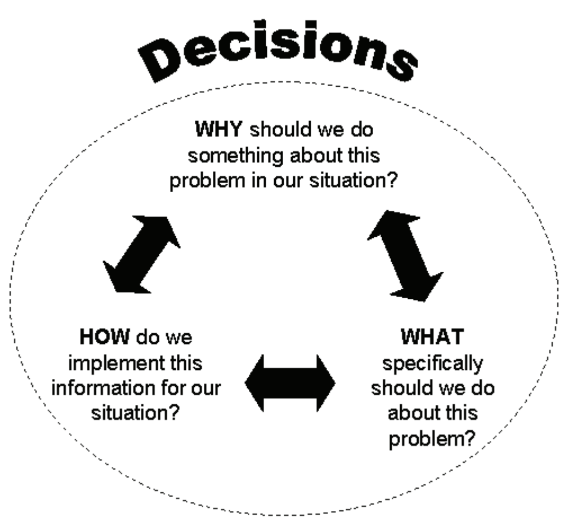


FIGURE 5-2 Questions that guide the gathering of evidence.

Some users of the framework may be tempted to skip this step, thinking they already know why they need to take action on obesity. However, the committee strongly recommends this as a first step in intervention planning. Table 5-1 lists some examples of areas of concern addressed by "Why" questions and the corresponding evidence that might be gathered.

"Why" questions help decision makers assess a public health problem in their particular region and locale. For example, a study characterizing determinants of obesity in an urban area identified individual and neighborhood contexts that were important considerations in formulating obesity-related policies (Black and Macinko, 2010).

To answer "Why" questions about obesity and other public health problems, decision makers require specific, often quantifiable evidence with which to evaluate the scope and severity of the problem, both in an absolute sense and relative to other issues requiring action. The evidence secured may help define characteristics of high-risk populations; it may also help develop benchmarks for setting goals and tracking progress. In cases where decisions have already been made, this kind of information is frequently needed to undergird and justify those decisions.

Studies that generate this type of evidence may include sample surveys or compilations of administrative data for monitoring and surveillance, population trend analysis, or studies of health impact, as well as studies that produce cost estimates and projections of future burden. For example, evidence that provides insight into obesity-related health care costs and lost productivity is helpful for comparing the eco-

TABLE 5-1 Areas of Concerns and Examples of Evidence Needed: "Why" Questions

Area of Concern	Examples of Evidence Needed
Public health situation	
• Health burden	Recorded levels of illness, disease, or death related to obesity
• Frequency/incidence of disease or risk factor	Number of people or rate of new cases affected by obesity or obesity-related diseases
• Social or environmental determinants of disease or risk factor	Number of catchment areas that do not have a supermarket or food stores offering healthful foods or suitable options for physical activity and exercise
• Trends	Rates of increase of obesity, obesity-related diseases, or adverse social determinants
• Health disparities	Relative or absolute differences in risk among demographic groups or subgroups
Monetary and social costs	
• Health care costs	Estimates of public dollars currently spent on providing health care for obesity or related conditions
• Other societal costs	Estimates of dollars spent by or lost from the public or private sector due to consequences of obesity (e.g., employee absenteeism)
	Estimates of other adverse outcomes associated with obesity with implications for social productivity

nomic burden of obesity with that associated with other health problems (Chenoweth and Associates, 2009; Finkelstein et al., 2009; Wolf and Colditz, 1998). Finkelstein and colleagues (2009), for instance, use data from the 1998 and 2006 Medical Expenditure Surveys to estimate health care costs attributable to obesity for all health care payers. When programs are competing for funding, such an evidence pool is necessary to support decisions.

"What" Questions

Locating evidence to answer "What" questions helps decision makers assess the effects of interventions with the potential to yield the best health outcomes and to ascertain the optimal conditions for obtaining those outcomes. Table 5-2 lists some examples of areas of concern addressed by "What" questions and the corresponding evidence that might be gathered.

To answer "What" questions, decision makers require evidence on the effects or impact of particular interventions on specific health outcomes over the short or long

TABLE 5-2 Areas of Concerns and Examples of Evidence Needed: "What" Questions

Area of Concern	Examples of Evidence Needed
Presumed mechanisms of intervention effects in target populations	
• Intervention theory or logic	Underlying assumptions (explicit or implicit) about how the intervention will improve health outcomes
• Causal pathways	Expected direct or indirect pathways linking the intervention to the outcomes at one or more levels and for different demographic groups
• Multiple causal levels for multiple influences	Intervening or interacting influences that might facilitate or hinder the effects in groups at risk
Effectiveness of intervention based on empirical studies or simulations	
• Links between intervention delivery and outcomes	Evidence that the intervention leads to the outcomes, including evidence of authentic and consistent implementation when effects were obtained
• Comparative outcomes	Effects of an intervention on the outcome in comparison with other intervention options or no intervention, including evidence on effects in different demographic groups
• Sustained effects	Evidence that effects of the appropriately implemented intervention are sustained over time
• Contextualized effects	Evidence of circumstances under which the evidence for effectiveness of the intervention is strongest, including evidence of other factors that influenced or interacted with the intervention (e.g., individual, family, community, or school factors)
• Unintended (unexpected) consequences	Positive or negative outcomes that lie outside the theory or underlying logic of the intervention

term. This is a core issue in the selection of an intervention or set of interventions. Decision makers need to understand how interventions work to change environments or behavior and whether they have additional indirect effects—positive or negative. A key aspect of "What" questions is whether a particular intervention is sufficient alone or requires other interventions at the same or different levels to have any effect or the maximum effect.

The most informative study designs for generating evidence of impact are comparative experiments or approximations of experiments that allow for evaluation of the effects of an intervention against a comparison condition or control group. The control group provides a reference point for what might have occurred in the absence of the intervention. Without a comparison condition of some type, one cannot determine whether changes observed were actually due to the intervention or might have occurred anyway as a result of other influences that coincided with the intervention. Outcomes may be compared against those of another intervention thought to have no effect on the problem being studied or with an alternative intervention designed to address the same problem. Comparisons may be for a population overall but may also focus on specific subgroups, for example, to assess whether the effects of the intervention and the alternative are similar in higher- versus lower-risk groups. Outcomes at one point in time may also be compared with outcomes assessed in the same population previously (e.g., a historical comparison or time series approach.

Studies that generate answers to "What" questions are referred to as impact assessments or outcome evaluations in the evaluation literature (Fitzpatrick et al., 2004; Rossi et al., 2004); they fall within the general category of effectiveness research in the social and behavioral sciences. The interventions evaluated might include programs, policies, laws, or some combination thereof operating at different levels of a region, community, organization, or institution. Kuo and colleagues (2009) provide a good example of a health impact assessment of a state law. Using published and unpublished data to model consumer response to point-of-purchase calorie postings at large chain restaurants, the authors quantify the potential impact of California's state menu labeling law on population weight gain in Los Angeles County.

Additional evidence of interest for "What" questions can be obtained by study designs that examine multiple pathways to outcomes (various causal mechanisms with direct and indirect effects), ripple effects (effects of an intervention on secondary outcomes that are linked to the main outcomes of interest), and unintended consequences (either positive or adverse effects that can be attributed to the intervention). For example, Schwartz and colleagues (2009) surveyed school students in Connecticut before and after low-nutrition snacks were removed from their schools to address concerns about potential unintended adverse effects (e.g., compensatory eating); these effects were not found after the intervention was implemented. A similar finding was reported from Arkansas based on surveys conducted after the statewide body mass index (BMI) screening and related school-based obesity prevention policies were implemented (Thompson and Card-Higginson, 2009) (see the discussion of this initiative later in this chapter).

"How" Questions

The potential effects of an intervention can be realized only if it is delivered appropriately in a particular setting. Answers to "How" questions help decision makers determine how to meet this need. Table 5-3 lists some examples of areas of concern addressed by "How" questions and the corresponding evidence that might be gathered.

To answer "How" questions, decision makers require evidence about the population, setting, and time frame at hand. Answers to "How" questions may feed back to "Why" and "What" questions by helping decision makers determine how plans and expectations should be adjusted to their context, what resources are needed to

TABLE 5-3 Areas of Concern and Examples of Evidence Needed: "How" Questions

Area of Concern	Examples of Evidence Needed
Relevance of the intervention on a large scale	
• Generalizability	Likelihood of achieving the expected outcomes in different demographic groups and in different communities/regions
• Sustainability	Likelihood that the intervention effect will last more than 1 to 5 years in the various groups
• Stakeholder acceptance	Likelihood that the intervention be well received by the target population/community and program delivery personnel
Costs and benefits of large-scale implementation	
• Cost-effectiveness	Costs of the intervention versus measured effects on the outcomes of interest (e.g., premature deaths averted, years of life saved, pounds lost) compared with other competitive interventions
• Cost/benefit	Value of health benefits versus the costs of implementation compared with other competitive interventions
• Cost feasibility	Costs of implementing the intervention on a communitywide scale compared with other competitive interventions
• Cost minimization	Costs of implementing the intervention in a hospital compared with costs of implementing a competitive intervention in a community clinic and achieving the same outcome
• Cost utility	Stakeholder preferences versus implementation costs compared with other competitive interventions
Political and practical concerns	
• Implementation priorities	Fit of the intervention with community or policy priorities; basis for giving this intervention high priority
• Portfolio balance	Fit within an overall set of interventions if considerations of feasibility, size of impact, and certainty of effect are combined
• Strategic planning	Strategies and tactics that can be used to mount this intervention
• Potential challenges	Implementation challenges that can be anticipated and evidence on how to overcome them

obtain the theoretical or desired effects, how interventions need to be adapted, what training is needed by staff, and how effects can be sustained over time. Answers to these questions relate to such concerns as feasibility, adaptability, and robustness of effects in natural contexts; acceptability to various stakeholder groups; replicability; sustainability; and resources needed for implementation. For example, the Pathways obesity prevention study in American Indian Children and the HEALTHY study to reduce risk factors for type 2 diabetes were school-based interventions that included extensive, detailed qualitative and quantitative measures to track implementation in addition to measures used to track study outcomes (Schneider et al., 2009; Steckler et al., 2003; Stone et al., 2003). Likewise, Wang and colleagues (2003) estimate the cost-effectiveness and cost/benefit of a completed school-based obesity prevention program (Planet Health), taking into account the costs of implementing the program as well as long-term outcomes, to assess the appropriateness of using public funds for this type of program. This form of evidence is often used to assess generalizability to other populations, settings, contexts, or time frames, referred to as "translation evidence" in public health and medicine.

Posing "Why," "What," and "How" Questions After a Policy or Program Is in Place

The above discussion addresses the types of concerns that guide information gathering before an intervention is undertaken. However, not all interventions relevant to obesity prevention have been based on a complete evidence picture. As noted earlier in this report, decision makers are not waiting for the best possible evidence but are acting on the available evidence because of the urgency of the problem. Initiatives may be debated in the public policy arena and proceed if the logic for their implementation is convincing and there is sufficient stakeholder and political support, but may subsequently be debated or criticized. Evidence is therefore required to document an intervention's implementation, specific outcomes, and further needs. This was the case in the Arkansas school-based child and adolescent obesity prevention initiative, whose components are listed in Box 5-1 (Ryan et al., 2006). These components were based on documentation of need and on national consensus recommendations for appropriate actions. After the initiative was launched, a program of annual monitoring of its implementation and outcomes was established. In such cases, the conceptual framework of "Why," "What," and "How" questions and the areas of concern identified in Tables 5-1 to 5-3 can readily be applied (Thompson and Card-Higginson, 2009).

Following implementation, "Why" questions used for needs assessment guide the gathering of information on whether the underlying situation has changed (e.g., using population data to quantify why an intervention is still needed or to justify funds that have been expended on the intervention by showing that progress has been made). In Arkansas, for example, standardized BMI screening data are used to track BMI changes and overall obesity prevalence and trends among schoolchildren (Thompson and Card-Higginson, 2009).

Box 5-1
Arkansas Framework for Combating Childhood and Adolescent Obesity, with National Recommendations for Action

Components of Arkansas activities	National Recommendations: U.S. Surgeon General	National Recommendations: AAP	National Recommendations: IOM
Assessment (primary data)			
BMI measurement and reporting		•	•
Individual BMI reports			
Subpopulation reports of obesity prevalence	•		•
School and school district reports of obesity prevalence			•
Statewide report of obesity prevalence	•		•
Pouring contract review disclosure			
School health index (CDC) in all schools			
Child Health Advisory Committee	•	•	•
Population Interventions			
Professional education requirements for cafeteria workers	•		•
Vending machine/competitive food restrictions and elimination from schools	•		•
Content requirements for vended/competitive foods	•		•
Physical activity requirements	•	•	•
Health education requirements	•	•	•
Local/district-level School Nutrition and Physical Activity Advisory Committees	•	•	•
Individual Interventions			
Child health reports to parents		•	•
Continuing medical education on obesity prevention/treatment for health professionals	•		•
Reimbursement plans for treating obesity	•	•	•
Surveillance			
Annual evaluation of Act 1220 programs (UAMS COPH)	•	•	•
Annual statewide public school assessment of BMI	•		•
Public health impact assessments by HHS/ADH			•

NOTE: AAP = American Academy of Pediatrics; ADH = Arkansas Department of Health; BMI = body mass index; CDC = Centers for Disease Control and Prevention; HHS = U.S. Department of Health and Human Services; IOM = Institute of Medicine; UAMS COPH = University of Arkansas for Medical Sciences College of Public Health.

SOURCE: Copyrighted and published by Project HOPE/Health Affairs as Ryan et al., 2006, "Arkansas fights fat: translating research into policy to combat childhood and adolescent obesity," Health Affairs 25(4), 992-1004. The published article is archived and available online at www.healthaffairs.org.

"What" questions, used to identify the effectiveness of interventions, guide the gathering of information on whether a specific intervention or interventions had the intended effects and contributed to addressing the problem in the expected manner or with the expected impact. In the Arkansas case, the subsequently established program of annual monitoring provides data that can be used to determine whether the various components of the initiative are resulting in improvements in the relevant school environment or child and parent behaviors. Likewise, definitive evidence for the effects on obesity prevention of such initiatives as posting calorie content on chain restaurant menus will not be available until the interventions have been in effect for a period of time. In the interim, effects on consumer behavior can be modeled for defined populations based on such data as the numbers of such restaurants, sales data on menu items with different calorie content, the types of consumers who patronize these restaurants, and how chain restaurant purchases influence their caloric intake.

"How" questions, used to identify issues related to relevance and implementation, can guide the gathering of information on what actually took place when an intervention was undertaken. Answers to these questions can be used to inform decisions about what changes might be needed going forward, to explain why the intervention might not have worked as well as intended, or to understand why costs were higher or lower than those initially estimated. The Arkansas monitoring program, for example, collects data on the quality and process of implementation (Thompson and Card-Higginson, 2009). Likewise, the methodology used in the Assessing Cost-Effectiveness in Obesity Study (ACE-Obesity) can help guide future decisions on interventions to prevent obesity in children and adolescents (Carter et al., 2009; Haby et al., 2006). In this study, researchers assessed absolute costs and potential cost savings to help determine the cost-effectiveness of 13 different interventions applied populationwide.

Applying the Evidence Typology: An Illustrative Example

In New York City, a decision was made to implement a policy requiring restaurants to publish calorie information on menus (Mello, 2009). "Why," "What," and "How" questions that guided, or may have guided, the search for and synthesis of evidence to support this broad-based policy decision are listed in Box 5-2. While the case report from which the questions were derived does not detail the exact decision-making processes, the relevant types of evidence that may have been used can be inferred from the documentation provided.

To approach this issue by applying the L.E.A.D. framework, decision makers would obtain a full picture of the scope and dimensions of the obesity problem in the city. For example, questions 1 to 5 in Box 5-2 could help decision makers assess the city's health needs comprehensively. Much of this evidence speaks to various aspects of the obesity prevalence issue, correlates of obesity, and relevant lifestyle factors in a situation-specific manner. This evidence might be collected by using archival data from

Box 5-2
Applying the Evidence Typology: A Case Study of New York City's Menu Labeling Policy

The Problem: "New York City is home to approximately 2 million overweight and 1 million obese adults. Diabetes has been diagnosed in nearly 10% of overweight adults and 18% of obese adults in the city, and an estimated 200,000 residents have undiagnosed diabetes. Hospitalization costs for New Yorkers with diabetes topped $481 million in 2003" (Mello, 2009, p. 2015).

Public Health Laws and Policy Actions to Address the Problem: Restaurants were required to make calorie information publicly available, posted on all menus and menu boards.

Types of Questions Decision Makers Might Ask to Grasp the Problem Comprehensively and Guide Evidence-Gathering Efforts:

"Why" take action?

1. What are the data on the prevalence of the problem? How widespread is the problem in different parts of the community and subpopulations?
2. What diseases are known to be associated with obesity, and what is the degree of association with serious diseases?
3. How many deaths result annually from diseases known to be related to obesity in New York City?
4. What is the cost burden of the city's obesity problem on the health care system?
5. On average, what is the annual frequency of visits to food service establishments and their use by the typical family in New York City? How many calories and what foods are purchased?

"What" action should we take that will give us the results we want?

6. What is the evidence that different types of community interventions (e.g., public education about calorie intake and obesity [which was the option chosen in New York City], community health campaigns, school-based programs) work in curbing the obesity problem (or a related public health problem) on a large scale?

"How" should we implement this action?

7. For the action chosen, what do experts tell us about the barriers to implementing a communitywide policy successfully? What means are effective for monitoring the policy implementation processes?
8. What is the parallel evidence on policy actions addressing other community health issues, such as smoking?

NOTE: This list is not intended to be exhaustive.

SOURCE: Mello, 2009.

hospital discharge records and/or survey data (see the previous discussion of "Why" questions), some of which could be used to examine relationships among relevant variables (e.g., caloric intake, obesity, and diabetes). Some of the evidence could be used in descriptive form. According to Mello (2009), this type of evidence gathering was actually conducted in the New York case.

Questions 6 to 8 in Box 5-2 would allow for a more complete problem assessment to examine potential impacts of candidate interventions. Causal evidence on particular interventions is important to answer "What" questions. Additionally, parallel practice evidence (that is, evidence about the effectiveness of other relevant interventions for other health or social issues) might be useful. Likewise, informed expert opinion could be a useful source of evidence for appraising what is likely to work before population-based investments are made. With limited resources, comparative cost evaluations might also be imperative.

A variety of sources, then, could be used to gather evidence applicable to this case, such as observational research, experimental and quasi-experimental studies, expert knowledge, and parallel evidence on the implementation of interventions. The next section elaborates on these sources.

POTENTIALLY USEFUL SOURCES OF EVIDENCE

Seven categories of study designs and sources of evidence may be useful for addressing the concerns listed in Tables 5-1 to 5-3:

- nonexperimental or observational studies,
- experimental and quasi-experimental studies,
- qualitative research and analysis,
- mixed-method studies,
- evidence synthesis methods,
- parallel evidence, and
- expert knowledge.

It is important to note that quality standards apply to all of these categories, and the relative value of each source depends on the decision-making context (see Chapter 6). Each category is discussed in turn below (see Appendix A for formal definitions and Appendix E for further discussion).

Nonexperimental or Observational Studies

An observational study is one in which the researcher assesses relationships among variables but does not manipulate an intervention or variables associated with potential outcomes. Study designs or methodologies that fall within this category include cross-sectional, longitudinal, or retrospective survey research or opinion polls, trend analysis, some types of market research, secondary analysis of existing databases, cohort and case-control studies, predictive studies, archival studies, census studies, monitoring and surveillance studies, implementation tracking, and policy analysis. These designs and methodologies are helpful for answering several "Why," "What," and "How" questions that can guide the search for associated evidence. Table 5-4 provides examples of how this type of evidence might be used. See Appendix E for a detailed discussion of economic cost analysis.

Cross-sectional studies, conducted at one point in time, raise the additional problem of identifying whether observed associations are temporal: Did a person become obese because of a low physical activity level, or did physical activity decline as weight increased, or both? Also, many cross-sectional surveys are conducted for general use or for specific administrative purposes and may vary in the completeness of coverage of relevant variables and the quality of data.

Despite these limitations, good-quality observational data are the best evidence sources for answering many questions of potential importance for decision making. They can also be useful as a source of baseline measures in populations.

Experimental and Quasi-experimental Studies

An experimental study is one in which the investigator has full control over the allocation of subjects to a preventive or treatment intervention versus a control condition, as well as the timing of an intervention. Randomized manipulation and assignment of individuals or groups to an intervention is a defining requirement of an experimental study. By contrast, a quasi-experimental study (e.g., matched cohort, regression-discontinuity, or interrupted time series design) is often described as nonrandomized because the investigator lacks full control over the allocation process and timing of the intervention. A quasi-experimental study design often includes pre–post intervention studies in which outcomes are measured both before and after the intervention is implemented.

Both experimental and quasi-experimental studies are potentially helpful sources of evidence for answering "What" questions about certain categories of interventions. For example, a quasi-natural experimental design was used to estimate the causal impact of physical education classes on overall student physical activity and weight (Cawley et al., 2007), and an interrupted time series design was used to evaluate the effectiveness of a framework designed to increase fruit and water consumption (Laurence et al., 2007). Table 5-5 provides examples of how this type of evidence might be tied to specific questions and applications.

As explained in Chapter 3, quasi-experiments are of critical importance as sources of evidence for obesity prevention interventions as an alternative to randomized controlled trials (RCTs). Experiments simply cannot be conducted with certain environmental or policy variables that influence obesity because they lie outside the control of researchers. Quasi-experimental approaches may be used for evaluating ongoing initiatives. See Appendix E for a detailed discussion of some commonly used

TABLE 5-4 Types of Observational Evidence and Examples of Their Uses

Type of Evidence	Questions That Could Be Addressed	Specific Applications
Quantitative Surveys, Longitudinal Studies, and Opinion Polls	Based on measurements or self-reports of height and weight, how many people are obese in a given region? (“Why” question)	A **cross-sectional sample survey** yielding descriptive data on levels of obesity using height and weight indicators
	Going forward, do self-reported eating choices change in a group of adolescents exposed to different levels of calorie information at their school cafeterias? (“What” question)	A **longitudinal cohort study** yielding analytic breakdowns of teens’ eating choices according to menu labeling policies in their respective school cafeterias
	Given two or more intervention options, which ones do stakeholders prefer? (“How” question)	A **cross-sectional sample survey** or poll yielding descriptive data on stakeholder preferences
Analysis of Existing Databases	Based on the health department’s data on a city population, what were the recorded levels of cardiovascular disease related to obesity in 2007? (“Why” question)	A **secondary analysis** of the **correlation or other measure of risk** between levels of measured obesity and cardiovascular disease in a selected **cross-sectional sample**
	Based on the health department’s data on a city population, what are the trends in cardiovascular disease and obesity levels from 2005 to 2010? (“Why” question)	A **trend analysis** comparing data for 2005 and 2010 for the percent of people with cardiovascular disease and the percent of people with obesity
	Based on the health department’s data on a city population, what are the past lifestyle correlates of cardiovascular disease in obese and nonobese adults in 2007? (“Why” question)	A **retrospective, case-control analysis** of the relationship between past lifestyle factors and cardiovascular disease in obese and nonobese adults
	Compared with another intervention, what are the costs for implementing and operating a school-based program districtwide, based on data from pilot or single-site studies? (“How” question)	A **cost-feasibility** analysis using preexisting budget and accounting databases
	What combination of factors maximally predicts stakeholder preferences with regard to program participation and use? (“How” question)	A **projection** analysis using preexisting databases on stakeholder preferences, participation, and program management
	What level of sales taxes or excise taxes on sugar-sweetened beverages directly results in decreased consumption of these beverages in a state by 50%? (“What” question)	A **modeling** study in which data on price elasticity, together with data on patterns of sales and consumption of these beverages, are used to estimate the effective level of taxation to decrease consumer consumption by 50%
	Which neighborhoods have the highest rates of childhood obesity, and what other characteristics of these neighborhoods might influence these rates? (“Why” question)	A **geographic mapping** study in which the locations of food stores and outdoor recreational facilities are plotted by neighborhoods or zones around neighborhoods alongside area data on child obesity prevalence
	What would be the potential reach of a policy to require calorie labeling on menus of chain restaurants? (“How” question)	Analysis of administrative data on food retail establishment locations and customer sales for the types of restaurants that would fall under the policy

TABLE 5-5 Types of Experimental and Quasi-Experimental Evidence and Examples of Their Uses

Type of Evidence	Questions That Could Be Addressed	Specific Applications
Randomized Controlled Trial (RCT)	Compared with another intervention, are the obesity outcomes better for individuals assigned at random to receive this intervention than for those assigned not to receive the intervention? ("What" question)	A **randomized controlled trial** of a manipulated nutrition program in two groups of obese adults who were placed in the program based on a coin flip (The random assignment usually balances individual characteristics across those who receive or do not receive the program so that the result can be interpreted as "all other things being equal"; some statistical controls may be required.)
Quasi-experimental Study	Compared with another intervention, are the obesity outcomes better with this intervention when administered to adults in two similar communities? ("What" question)	A **matched-cohort study** design comparing obesity outcomes of a manipulated nutrition program in two communities (The two communities are the groups that are matched on relevant characteristics; other potential influences on intervention outcomes are statistically controlled.)
	Using ongoing obesity measures as control data in a group of children, is body mass reduced when this intervention is administered in alternating cycles? ("What" question)	An **interrupted time series study** tracking changes in obesity outcomes over time when a nutrition program is administered periodically

quasi-experimental approaches in other fields that are recommended for increased emphasis in research on obesity prevention.

Qualitative Research

Qualitative research evidence is typically derived from documentary sources, field observations, interviews, and open-ended verbal interactions between participants and researchers. Examples of studies that fall within this category include logic modeling or program theory analysis, ethnographic studies, focus group or key informant interviews, content or documentary analysis, case studies, some intervention process delivery and implementation monitoring, and evaluability assessments of programs and interventions. Qualitative research approaches give researchers the ability to assess perceptions of respondents at a much richer level than is possible using questionnaires with fixed responses.

Qualitative methods employ emergent designs (not preset, as in quantitative methods) and contextualized understandings of phenomena. Those investigating an intervention will want in-depth information on how people interact with the intervention and with different variables in that context. Questions and data-gathering techniques may be expanded or modified as data are collected. This type of research may be helpful in answering all three types of questions; Table 5-6 provides examples illustrating how this type of evidence can address "What" and "How" questions.

Limitations of qualitative methods lie in the subjectivity they introduce by making the researcher an instrument of the research process and the difficulty of bringing closure to open-ended forms of inquiry. There are well-developed conventions for adding rigor to qualitative inquiry, such as triangulation, convergent validation, and internal and external criticism (see Chapter 6).

Mixed-Method Studies

Mixed-method studies employ methodologies drawn from a variety of disciplines, including both qualitative and quantitative data gathering and analysis methods. Studies may combine extensive descriptions of context and the experiences of program participants with standardized assessments of changes in institutional, environmental, or individual behavior–related variables. The realization that these types of data are complementary has increased interest in the use of such studies in the public health arena. Examples of mixed-method studies include surveys and interviews combined with RCTs, interviews combined with interrupted time series analysis, policy-related content analysis combined with focus group interviews, health impact analysis using archival databases and surveys, economic analysis using archival databases and surveys, systems mapping based on a review of the literature, simulation studies, and mixed-method evidence synthesis techniques (discussed in the next section). Roux and colleagues (2008) use a combination of a systematic review of disease burden and data from clinical trials, population-based surveys, and other published literature to assess the cost-effectiveness of community-based physical activity interventions associated with disease incidence. This type of research is particularly helpful in answering "What" questions. Mixed-method studies may also help in garnering evidence to

TABLE 5-6 Types of Qualitative Evidence and Examples of Their Uses

Type of Evidence	Questions That Can Be Addressed	Specific Applications
Logic Modeling or Program Theory Analysis	What are the underlying assumptions about how an intervention will improve health outcomes? What are the expected causal pathways? What intervening factors in the larger system and community are likely to affect outcomes? ("What" questions)	A **content analysis** and **systematic review** of documents and literature relevant to an intervention to develop a logic model or causal path diagram
Process Delivery and Implementation Monitoring	What features of program implementation are associated with the maximum effect of this program? ("What" question) What are the documented barriers to implementation of this intervention, and how have they been overcome? ("How" question)	A **qualitative focus group interview** of program delivery personnel from an effective program

address "Why" and "How" questions. For example, Manning and colleagues (1991) assess the economic costs of poor health behaviors, including a sedentary lifestyle. This information can be helpful in addressing why something should be done to increase the level of physical activity in a community. Using data from the National Health Interview Survey and the RAND Health Insurance Experiment, the authors employ an incidence-based approach to determine the lifetime costs of a sedentary lifestyle. Table 5-7 provides examples of the uses of this type of evidence.

Evidence Synthesis Methods

Evidence synthesis methods encompass systematic reviews, and meta-analyses of experimental and/or quasi-experimental studies, syntheses of qualitative research, and mixed-method evidence syntheses. Evidence synthesis is particularly useful in answering "What" questions, but may also be useful for assembling information to address "Why" and "How" questions. Table 5-8 provides examples of how this type of evidence might be used.

Systematic Reviews

The literature on evidence synthesis distinguishes between *primary studies*, or individual studies presenting original data, and *systematic reviews*, or organized summaries of a body of research that addresses a focused question using methods intended to reduce the likelihood of bias. Traditional methods of evidence synthesis (in the evidence-based medicine tradition) have relied heavily on quantitative methods, such as meta-analysis, to generate summaries of research-based evidence (see Hunt, 1997; Smith and Glass, 1977). Several systematic reviews relevant to obesity prevention are listed in Appendix C. Some of the limitations of this type of evidence synthesis when applied to public health issues are discussed in Chapter 3.

TABLE 5-7 Types of Mixed-Method Evidence and Examples of Their Uses

Type of Evidence	Questions That Can Be Addressed	Specific Applications
Surveys Combined with Randomized Controlled Trial (RCT)	What self-reported individual-, family-, and community-level factors moderate effects on obesity outcomes when an intervention is randomly assigned to one of two adult groups? ("What" question)	A **randomized controlled trial** of a manipulated nutrition program using two groups of obese adults, with **survey-based data** of individual-, family-, and community-level factors **analyzed as potential moderators of** outcomes
Qualitative Analysis Combined with Quasi-experimental Study	Is there qualitative evidence to show that an intervention was implemented as intended when outcomes improved in a time series analysis? How consistently and authentically was the intervention implemented when effects were obtained? ("What" questions)	A **content analysis** of food logs, menus, and meal plans combined with an **interrupted time series study** tracking changes in obesity outcomes when a nutrition program is implemented periodically

TABLE 5-8 Types of Evidence Synthesis Methods and Examples of Their Uses

Type of Evidence	Questions That Can Be Addressed	Specific Applications
Systematic Reviews: Experimental and/or Quasi-experimental Studies	Based on formal syntheses of experimental or quasi-experimental research, what is the evidence on the effectiveness of this intervention? ("What" question)	A **systematic review** of effects of mandatory school exercise programs on childhood obesity levels
Meta-analyses: Experimental and/or Quasi-experimental Studies	Based on meta-analysis of effects from experimental or quasi-experimental research, what is the evidence on the effectiveness of this intervention? ("What" question)	A **meta-analysis** to estimate the average effect on childhood obesity levels found in eligible studies of mandatory school exercise programs
Mixed-Method Syntheses: Stakeholder Studies	Based on a formal summary of results, what are the facilitators of and barriers to implementation of this intervention in light of stakeholder perspectives? ("What" question)	A **"realist" review** using mixed-method analysis of stakeholder participation studies (drawing on the realist philosophy of science)

Meta-Analyses

Meta-analysis is a statistical procedure that pools results from a sample of preexisting experimental/quasi-experimental studies to derive a single *effect size*. Effect size is a quantitative index expressing the difference between the treatment and control group means in standard deviation units. Studies are selected on the basis of some preset criteria, typically a common question about an intervention or therapeutic procedure, a defined target population, and common outcomes of interest.

The chief difficulty with applying meta-analysis is that the number of available studies on a given topic is usually rather small. This difficulty forces researchers to use a backwards logic in justifying the procedure—to start with a selected sample of studies and then imagine a hypothetical population in which the studies belong. Given this scenario, assumptions can never be properly tested; Glass (2000) finds this logic to be untenable today. Further, the studies selected typically vary in the effects they show, from positive to negative. Technical advances now make it possible to improve the analysis, for example, by testing for and statistically ruling out heterogeneity in the sample (Cooper and Hedges, 1994). However, Glass (2000) finds these added procedures to be problematic, as further assumptions need to be made that are both untestable and, often, not defensible.

The most important limitation, according to Glass (2000), is that the meta-analysis is guided by a very limiting question: On average, is the (intervention) effective? The procedure cannot address questions about differential effects of an intervention, reducing findings across studies to an average that frequently removes the most important information about the intervention. It is this loss of information through averaging that Glass now finds regrettable. His current view is that the standard guiding question in meta-analysis needs to be replaced with deeper and broader questions, such as "What type of therapy, with what type of client, produces what kind of effect?" (Glass, 2000). Looking to the future, he suggests that "meta-analysis needs to be replaced by archives of raw data that permit the construction of complex data landscapes that depict the relationships among independent, dependent and mediating variables" (Glass, 2000).

Syntheses of Qualitative Research

Despite some initial controversy, a place for the inclusion of qualitative research in systematic reviews and evidence syntheses was established in medicine in 2001 in an editorial in the *British Medical Journal* (Dixon-Woods and Fitzpatrick, 2001). Citing Cochrane reviews that could have benefited from the use of qualitative research to assess how children and adolescents with cancer communicate about their disease, the editors called for expanded inclusion criteria incorporating relevant types of qualitative evidence to answer complex questions involving social variables and outcomes. At the same time, they cautioned evidence users about the need for rigor in identifying and appraising the quality of research evidence. The methodological challenges of distilling information from qualitative studies—frequently representing diverse theoretical, disciplinary, and analytic perspectives—into secondary summaries should not be underestimated (Dixon-Woods et al., 2001).

Likewise, Bower and Scambler (2007) demonstrate the relevance and utility of qualitative data in expanding the scope of evidence-based practice in dental public health. In their view, methods should be aligned with questions asked (see also Petticrew and Roberts, 2003).

Mixed-Method Evidence Synthesis

Qualitative and quantitative methods play complementary but contrasting roles in facilitating different types of inquiry aimed at better understanding practice-related phenomena (Bower and Scambler, 2007, p. 161). Traditional quantitative methods address questions about the prevalence of a disease or disorder, the statistical significance of an effect, the strongest predictors of a condition, program, or policy impact, or cost considerations in numerical terms. In contrast, qualitative methods help researchers ask questions about the nature, content, process or meaning of a particular process or outcome. Mixed-method approaches are suitable when questions cross these boundaries. One example in Bower and Scambler (2007) deals with examining the impact of advertisements on dental care and prevention of caries. In this case, social and personal factors outside an intervention, documented using both quantitative and qualitative methods, were found to mediate health-seeking behaviors of individuals and affected the expected outcomes.

Newer methods for synthesis of findings from qualitative studies and evidence from trials or other quantitative studies are now available in the fields of obesity prevention (Connelly et al., 2007; Thomas et al., 2004), medicine (Barbour and Barbour, 2003; Dixon-Woods et al., 2001), nursing (McInnes and Chambers, 2008; McNaughton, 2000), and public health. Following is an example that shows how evidence screening and data compilation and analysis can be approached by synthesizing diverse types and sources of evidence.

In a mixed-method synthesis of evidence, Connelly and colleagues (2007) combined a systematic review of eligible quantitative trials with qualitative content analysis of descriptive data on interventions, as reported in the studies. Because the approach relies on the realist philosophy of science that examines causal pathways while acknowledging social realities (as opposed to the logical-positivistic philosophy of traditional laboratory science), this approach to evidence synthesis has been referred to as a "realist review." The topic was prevention of childhood obesity and overweight. The authors began with a search of four electronic databases (including MEDLINE and PsychInfo) and applied published criteria to screen studies for quality and inclusion (e.g., execution of random assignment or other controls in study designs, sample size, lack of subject attrition, measurement validity and reliability, treatment duration and fidelity). This procedure initially identified 30 RCTs or controlled trials. Trials were included if they had a measured index of adiposity (the outcome), a population aged 0-18, and a 12-week follow-up to treatment. Blind to the study outcomes, the authors then sorted the studies into two categories based on a systematic content analysis of the intervention descriptions provided, using an "intensity of physical activity" indicator. This procedure allowed them to qualitatively separate "compulsory" from "voluntary" activity interventions. The authors then proceeded with the evidence synthesis, summarizing effects found across studies by group. Their conclusions showed that while nutrition education, skills training, and physical education were not effective in controlling childhood obesity, compulsory rather than voluntary physical activity was. Notably, the effects were not evident with the undifferentiated physical activity indicator in the overall sample. According to the authors, their rational, realist review permitted them to link the physical activity type causally with decreases in overweight and obesity.

Parallel Evidence

Parallel practice evidence is an indirect form of research evidence. Evidence of the effectiveness of an intervention in addressing another public health issue using similar strategies—such as the role of social marketing, curriculum programs, or financial factors in reducing smoking, speeding, or sun exposure—may have implications for an obesity-related outcome such as dietary intake. Here, the characteristics of the original research or evaluation reports, such as how the evidence was generated, whether the work was peer-reviewed, where the work is published, and the credentials of the author(s), help determine the quality of the evidence (Swinburn et al., 2005). Should such data be available in the form of a compiled policy brief or legal brief, again the quality of the evidence should be carefully evaluated based on the original data sources. Ideally, information about quality assessment would be incorporated into the report, with an explanation of the criteria used and a clear statement of the study limitations. This does not preclude checking original sources where feasible, but helps in judging the credibility of secondary sources. In general, one should also always rule out any conflicts of interest in secondary compilations of evidence. Parallel evidence is particularly helpful in answering "What" and "How" questions. Table 5-9 provides examples of how this type of evidence might be used. Also see Chapter 2 for examples of parallel evidence, including activities surrounding tobacco and alcohol use and HIV/AIDS.

Expert Knowledge

Expert knowledge includes the views of professionals with expertise in a particular field of practice or inquiry, such as clinical practice or research methodology. Expert opinion may refer to one person's views or to the consensus view of a group of experts. Consensus views are sometimes based on systematic reviews or other forms of evidence synthesis but still require interpretation or judgment about the evidence that is collected, and may require drawing conclusions in the absence of any or enough relevant evidence. Expert consensus and expert opinion should be appraised based on the credentials and experience of the experts involved, supporting documentation, the transparency and rigor of the consensus process, and ruling out of any conflicts

TABLE 5-9 Types of Parallel Evidence and Examples of Their Uses

Type of Evidence	Questions That Can Be Addressed	Specific Applications
Research Evidence on Effects of Parallel Interventions	Given the existing evidence on the effectiveness of tobacco and alcohol taxes, would soda taxes be similarly effective in reducing obesity on a large scale? ("What" question)	**Intervention impact or effectiveness studies** showing that strategies to influence public behaviors work
Parallel Research on Legal Issues	Given the constitutional issues involved in restricting free speech, what grounds have been used to justify controls on advertising? ("What" question)	**Content analysis** of relevant cases to identify arguments that have been advanced for and against such restrictions and how these arguments have been resolved
Parallel Research on Implementation Process or Policy Development	Given that eating and physical activity are individual behaviors but are affected by policies and programs in the broader community, what are some precedents for environmental and policy approaches that impact personal behavior, and how were they achieved and justified? ("What" question)	**Case studies** of effects of populationwide interventions on nutrition, physical activity, obesity, or cardiovascular disease or obesity in other countries[a]
	Given that obesity is a populationwide problem for which many of the drivers are a part of the social fabric, what can be learned from approaches used in other public health efforts of similar scope and complexity? ("How" question)	**Retrospective case analysis** of successes related to other complex public health issues that have been addressed by policies that have led to social change[a]

[a] See Economos et al., 2001; Eriksen, 2005; Kersh and Morone, 2002.

of interest. For example, participation and selection processes for committees that generate reports such as the *Guide to Community Preventive Services* (http://www.thecommunityguide.org/) or Common Community Measures for Obesity Prevention (http://www.activelivingresearch.org/node/11698) are highly rigorous and systematic. Expert knowledge is particularly helpful in answering "What" or "How" questions, although consensus reports usually also include a rationale for why action is needed. Examples of this type of evidence include national committee reports based on deliberative processes; guidelines from national associations, health foundations, and committed practitioners or health professional organizations; and other expert statements. Table 5-10 provides examples of this type of evidence and how it might be used. Also see Chapter 7 for further discussion of the use of expert opinion and experience.

GATHERING THE EVIDENCE

Locating evidence from the expanded perspective proposed by the committee requires access to a broad base of resources that extends beyond research journals. Databases used by researchers and decision makers and their intermediaries typically incorporate

TABLE 5-10 Types of Expert Knowledge and Examples of Their Uses

Type of Evidence	Questions That Can Be Addressed	Specific Applications
National Committee Reports Based on a Deliberative Process	In view of present and potential human and monetary costs of treating current levels of obesity or obesity-related diseases, what types of actions are recommended within and outside of the health care system? ("What" question)	Reasoned and formal **analysis by a committee of established experts**, such as a report by a committee convened by the public or private sector with established bias and conflict-of-interest procedures
Guidelines from National Associations, Health Foundations, and Committed Practitioner or Health Professional Organizations	Given the availability of effective measures to treat high blood pressure, what steps are needed to improve levels of blood pressure control in the population at large? ("What" question)	Public **statements of consensus** by a committee of established professionals and practitioners
Other Expert Statements	How does this intervention fit with community politics or national policy priorities? On what basis should it be given high priority? ("How" questions)	The **considered opinion** of experts in a particular field or practitioners, leaders, stakeholders, and policy makers able to make informed judgments on implementation issues and having local or governmental expertise (e.g., doctors, lawyers, nutritionists, scientists, or academics able to interpret the scientific literature or specialized forms of data, such as legal evidence)

data, studies, or reports from numerous disciplines but may still miss many potentially useful sources, such as compilations from economics, education, business, and law and information from newspapers, government documents, and reports of community agencies and programs. Appendix D provides guidance on where to gather evidence. This aspect of the process is presumed to rely primarily on electronic databases. Some relevant databases are in the public domain or accessible in public facilities; others require subscriptions or access through institutions such as university libraries. The search process generally involves identification of a set of key concepts and terms that can be used as probes in these databases. A preliminary screening of titles and abstracts of identified items for relevance is followed by detailed examination of items selected for potential use. Some types of evidence can be readily identified and evaluated by lay persons or professionals without scientific expertise. Examples include program descriptions, news reports, "report cards," and other compilations or analyses intended for direct use by lay audiences or nonresearchers. However, identifying and sifting through research studies requires academic or scientific expertise.

Awareness of the wide array of evidence sources and potentially relevant content areas is critical for full implementation of the L.E.A.D. framework. The value of electronic databases is in providing easily searchable centralized compilations of information within specific fields of knowledge. Unfortunately, however, although some databases overlap and some provisions exist for searching multiple databases simultaneously, there are gaps related to boundaries between disciplines. For example, those searching biomedical or health databases may encounter a limited amount of information from management science, law, public policy, marketing, education, communications, or anthropology—although all of these fields include both content and methodologies that are relevant to aspects of obesity prevention and public health in general. Also of potential relevance are electronic compilations of newspaper articles, program reports and evaluations, scientific and technical reports, and public surveys that are not available in academic journals. The listing in Appendix D includes databases for these sources as well as the academic literature.

Evidence sources should be evaluated to determine their degree of relevance to the user's question and context based on the nature of the outcomes for which the evidence is informative. Evidence quality should be judged by standards appropriate to that type of evidence. The evidence evaluation step in the L.E.A.D. framework is discussed in Chapter 6. Evidence gaps will be identified in this process and in some cases may predominate. That is, the available evidence—even when broadly sought—may not address decision makers' questions because no studies relevant to the issue or to the potential setting have been done, or because important studies do not report key details needed to fully interpret and apply the results from a practical perspective. Transmitting information about what evidence is not available to researchers is critical to the process of generating evidence to fill these gaps. Approaches to using actual ongoing policy and programmatic initiatives as the basis for generating evidence are discussed in Chapter 8.

REFERENCES

Barbour, R. S., and M. Barbour. 2003. Evaluating and synthesizing qualitative research: The need to develop a distinctive approach. *Journal of Evaluation in Clinical Practice* 9(2):179-186.

Black, J. L., and J. Macinko. 2010. The changing distribution and determinants of obesity in the neighborhoods of New York City, 2003-2007. *American Journal of Epidemiology* 171(7):765-775.

Bower, E., and S. Scambler. 2007. The contributions of qualitative research towards dental public health practice. *Community Dentistry and Oral Epidemiology* 35(3):161-169.

Carter, R., M. Moodie, A. Markwick, A. Magnus, T. Vos, B. Swinburn, and M. M. Haby. 2009. Assessing cost-effectiveness in obesity (ACE-obesity): An overview of the ACE approach, economic methods and cost results. *BMC Public Health* 9(Article 419).

Cawley, J., C. Meyerhoefer, and D. Newhouse. 2007. The impact of state physical education requirements on youth physical activity and overweight. *Health Economics* 16(12):1287-1301.

Chenoweth and Associates. 2009. *The economic costs of overweight, obesity, and physical inactivity among California adults—2006*. New Bern, NC: Chenoweth and Associates, Inc.

Connelly, J. B., M. J. Duaso, and G. Butler. 2007. A systematic review of controlled trials of interventions to prevent childhood obesity and overweight: A realistic synthesis of the evidence. *Public Health* 121(7):510-517.

Cooper, H., and L. V. Hedges. 1994. *The handbook of research synthesis*. New York: The Russell Sage Foundation.

Dixon-Woods, M., and R. Fitzpatrick. 2001. Qualitative research in systematic reviews. *British Medical Journal* 323(7316):765-766.

Dixon-Woods, M., R. Fitzpatrick, and K. Roberts. 2001. Including qualitative research in systematic reviews: Opportunities and problems. *Journal of Evaluation in Clinical Practice* 7(2):125-133.

Economos, C. D., R. C. Brownson, M. A. DeAngelis, S. B. Foerster, C. Tucker Foreman, J. Gregson, S. K. Kumanyika, and R. R. Pate. 2001. What lessons have been learned from other attempts to guide social change? *Nutrition Reviews* 59(3, Supplement 2):S40-S46.

Eriksen, M. 2005. Lessons learned from public health efforts and their relevance to preventing childhood obesity. In *Preventing childhood obesity: Health in the balance*, edited by J. Koplan, C. Liverman, and V. Kraak. Washington, DC: The National Academies Press. Pp. 343-375.

Finkelstein, E. A., J. G. Trogdon, J. W. Cohen, and W. Dietz. 2009. Annual medical spending attributable to obesity: Payer and service-specific estimates. *Health Affairs* 28(5): w822-w831.

Fitzpatrick, J. L., J. R. Sanders, and B. R. Worthen. 2004. *Program evaluation: Alternative approaches and practical guidelines*. 3rd ed. Boston, MA: Allyn and Bacon.

Glass, G. 2000. *Meta-analysis at 25*. http://glass.ed.asu.edu/gene/papers/meta25.html (accessed December 24, 2008).

Haby, M. M., T. Vos, R. Carter, M. Moodie, A. Markwick, A. Magnus, K. S. Tay-Teo, and B. Swinburn. 2006. A new approach to assessing the health benefit from obesity interventions in children and adolescents: The assessing cost-effectiveness in obesity project. *International Journal of Obesity (London)* 30(10):1463-1475.

Hunt, M. 1997. *How science takes stock: The story of meta-analysis*. New York: Russel Sage Foundation.

Kersh, R., and J. Morone. 2002. The politics of obesity: Seven steps to government action. *Health Affairs* 21(6):142-153.

Kuo, T., C. J. Jarosz, P. Simon, and J. E. Fielding. 2009. Menu labeling as a potential strategy for combating the obesity epidemic: A health impact assessment. *American Journal of Public Health* 99(9):1680-1686.

Laurence, S., R. Peterken, and C. Burns. 2007. Fresh Kids: the efficacy of a Health Promoting Schools approach to increasing consumption of fruit and water in Australia. *Health Promotion International* 22(3):218-226.

Manning, W. G., E. B. Keeler, J. P. Newhouse, E. M. Sloss, and J. Wasserman. 1991. *The costs of poor health habits*. Cambridge, MA: Harvard University Press.

McInnes, R. J., and J. A. Chambers. 2008. Supporting breastfeeding mothers: Qualitative synthesis. *Journal of Advanced Nursing* 62(4):407-427.

McNaughton, D. B. 2000. A synthesis of qualitative home visiting research. *Public Health Nursing* 17(6):405-414.

Mello, M. M. 2009. New York City's war on fat. *New England Journal of Medicine* 360(19):2015-2020.

Petticrew, M., and H. Roberts. 2003. Evidence, hierarchies, and typologies: Horses for courses. *Journal of Epidemiology and Community Health* 57(7):527-529.

Rossi, P. H., M. W. Lipsey, and H. E. Freeman. 2004. *Evaluation: A systematic approach*. 7th ed. Thousand Oaks, CA: Sage Publications.

Roux, L., M. Pratt, T. O. Tengs, M. M. Yore, T. L. Yanagawa, J. Van Den Bos, C. Rutt, R. C. Brownson, K. E. Powell, G. Heath, H. W. Kohl, 3rd, S. Teutsch, J. Cawley, I. M. Lee, L. West, and D. M. Buchner. 2008. Cost effectiveness of community-based physical activity interventions. *American Journal of Preventive Medicine* 35(6):578-588.

Ryan, K. W., P. Card-Higginson, S. G. McCarthy, M. B. Justus, and J. W. Thompson. 2006. Arkansas fights fat: Translating research into policy to combat childhood and adolescent obesity. *Health Affairs* 25(4):992-1004.

Schneider, M., W. J. Hall, A. E. Hernandez, K. Hindes, G. Montez, T. Pham, L. Rosen, A. Sleigh, D. Thompson, S. L. Volpe, A. Zeveloff, and A. Steckler. 2009. Rationale, design and methods for process evaluation in the HEALTHY study. *International Journal of Obesity* 33(Supplement 4):S60-S67.

Schwartz, M. B., S. A. Novak, and S. S. Fiore. 2009. The impact of removing snacks of low nutritional value from middle schools. *Health Education Behavior* 36(6):999-1011.

Smith, M. L., and G. V. Glass. 1977. Meta-analysis of psychotherapy outcome studies. *American Psychologist* 32(9):752-760.

Steckler, A., B. Ethelbah, C. J. Martin, D. Stewart, M. Pardilla, J. Gittelsohn, E. Stone, D. Fenn, M. Smyth, and M. Vu. 2003. Pathways process evaluation results: A school-based prevention trial to promote healthful diet and physical activity in American Indian third, fourth, and fifth grade students. *Preventive Medicine* 37(Supplement 1):S80-S90.

Stone, E. J., J. E. Norman, S. M. Davis, D. Stewart, T. E. Clay, B. Caballero, T. G. Lohman, and D. M. Murray. 2003. Design, implementation, and quality control in the Pathways American-Indian multicenter trial. *Preventive Medicine* 37(Supplement 1):S13-S23.

Swinburn, B., T. Gill, and S. Kumanyika. 2005. Obesity prevention: A proposed framework for translating evidence into action. *Obesity Reviews* 6(1):23-33.

Thomas, J., A. Harden, A. Oakley, S. Oliver, K. Sutcliffe, R. Rees, G. Brunton, and J. Kavanagh. 2004. Integrating qualitative research with trials in systematic reviews. *British Medical Journal* 328(7446):1010-1012.

Thompson, J. W., and P. Card-Higginson. 2009. Arkansas' experience: Statewide surveillance and parental information on the child obesity epidemic. *Pediatrics* 124(Supplement 1): S73-S82.

Wang, L. Y., Q. Yang, R. Lowry, and H. Wechsler. 2003. Economic analysis of a school-based obesity prevention program. *Obesity Research* 11(11):1313-1324.

Wolf, A. M., and G. A. Colditz. 1998. Current estimates of the economic cost of obesity in the United States. *Obesity Research* 6(2):97-106.

6

Evaluating Evidence

KEY MESSAGES

- What constitutes the best-quality evidence varies with the question being asked, which should be aligned with user needs and interests.
- Evidence should be evaluated against outcomes that may be short-term, intermediate, or long-term. These outcomes may be related to systems considerations (see Chapter 4) and not directly to the problem of interest.
- In evaluating the evidence for an intervention, both the level of certainty of the causal relationship between the intervention and its observed outcomes and the generalizability of the evidence to other individuals, settings, contexts, and time frames should be considered.
- The quality of each type of evidence needed to answer a particular question should be evaluated according to established criteria for that type of evidence.
- Evaluating the evidence gathered to address a particular population-level health problem will help identify gaps in knowledge that require further research.

The previous chapter describes an expanded perspective on the types of evidence that can be used in decision making for interventions addressing obesity and other complex, systems-level population health problems. It presents a detailed typology of evidence that goes beyond the traditional simple evidence hierarchies that have been used in clinical practice and less complex public health interventions. This chapter focuses on the question of how one judges the quality of different types of evidence in making decisions about what interventions to undertake. The question is an important one not only because many of the interventions required to address obesity are complex, but also because the available evidence for such interventions comes from studies and program evaluations that often are purposely excluded from systematic reviews and practice guidelines, in which studies are selected on the basis of the conventional hierarchies.

In the L.E.A.D. framework (Figure 6-1), one begins with a practical question to be answered rather than a theory to be tested or a particular study design (Green and Kreuter, 2005; Sackett and Wennberg, 1997). A decision maker, say, a busy health department director or staff member, will have recognized a certain problem or opportunity and asked, "What should I do?" or "What is our status on this issue?" Either of these questions may be of interest only to this decision maker for the particular social, cultural, political, economic, and physical context in which he/she works, and the answer may have limited generalizability. This lack of generalizability may lead some in the academic community to value such evidence less than that from randomized controlled trials (RCTs). However, data that are contextually relevant to one setting are often more, not less, relevant and useful to decision makers in other settings than highly controlled trial data drawn from unrepresentative samples of unrepresentative populations, with highly trained personnel conducting the interventions under tightly supervised protocols (see Chapter 3 for further discussion).

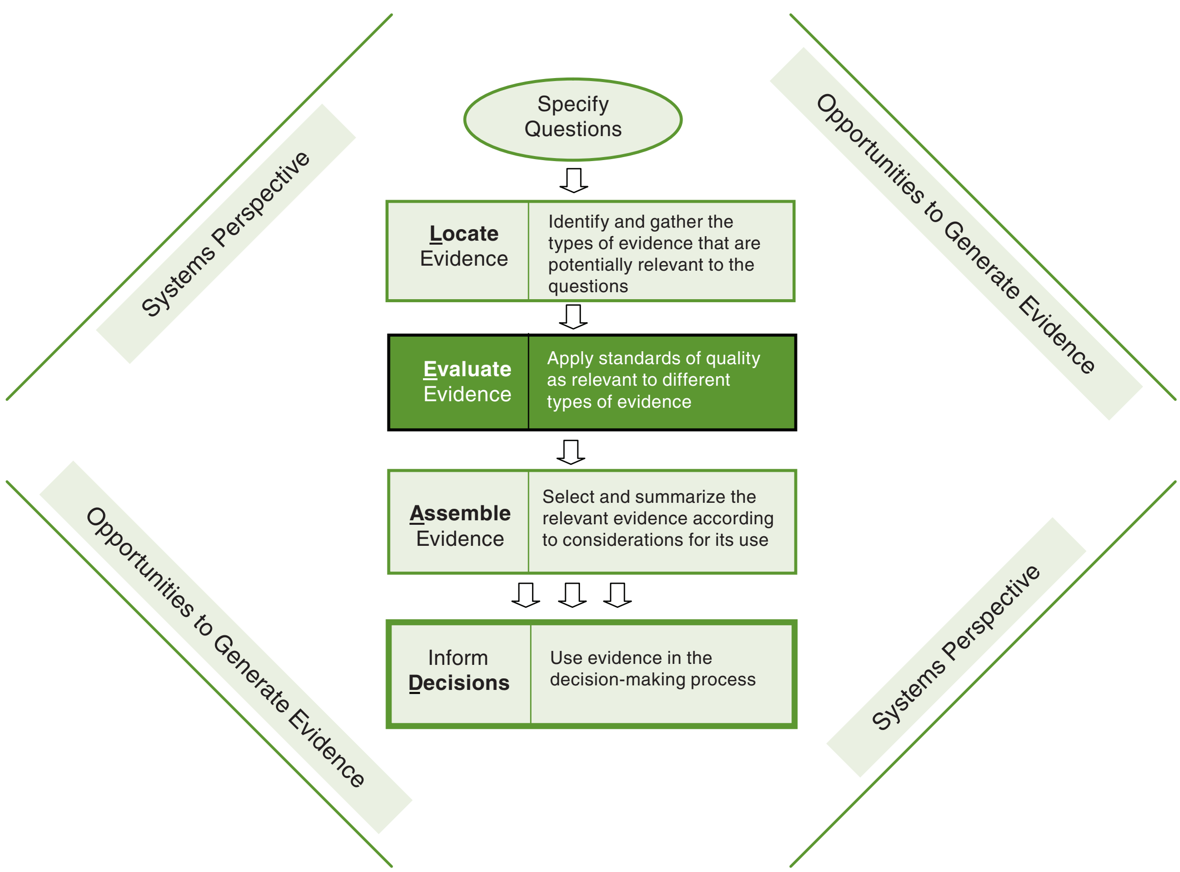


FIGURE 6-1 The **L**ocate Evidence, **E**valuate Evidence, **A**ssemble Evidence, Inform **D**ecisions (L.E.A.D.) framework for obesity prevention decision making.
NOTE: The element of the framework addressed in this chapter is highlighted.

The types of evidence that are used in local decision making, including the policy process, extend beyond research to encompass politics, economics, stakeholder ideas and interests, and general knowledge and information (see Chapter 3), and the decision maker needs to take a practical approach to incorporating this evidence into real-life challenges. Working from this expanded view of what constitutes relevant evidence and where to find it (Chapter 5), this chapter describes an approach for evaluating these different types of evidence that is dependent on the question being asked and the context in which it arises.

Before proceeding, it is worth emphasizing that the L.E.A.D. framework is useful not only for decision makers and their intermediaries but also for those who generate evidence (e.g., scientists, researchers, funders, publishers), a point captured by the phrase "opportunities to generate evidence" surrounding the steps in the framework (Figure 6-1). In fact, a key premise of the L.E.A.D. framework is that research generators need to give higher priority to the needs of decision makers in their research designs and data collection efforts. To this end, the use of the framework and the evaluation of evidence in the appropriate context will identify gaps in knowledge that require further investigation and research.

This chapter begins by reviewing several key aspects of the evaluation of evidence: the importance of the user perspective, the need to identify appropriate outcomes, and the essential role of generalizability and contextual considerations. After summarizing existing approaches to evaluating the quality of evidence, the chapter describes the general approach proposed by the committee. Finally, the chapter addresses the issue of the trade-offs that have to be made when the available evidence has limitations for answering the question(s) at hand—a particular concern for those who must make decisions about complex, multilevel public health interventions such as obesity prevention.

A USER'S PERSPECTIVE

The approach of "horses for courses" (Petticrew and Roberts, 2003) emphasizes that what constitutes best evidence varies with the question being addressed and that there is no value in forcing the same type of evidence to fit all uses. Once the question being asked is clear, users of the L.E.A.D. framework must either search for or generate (see Chapter 8) the kinds of evidence that will be helpful in answering that question. The next chapter describes how to assemble the evidence to inform decisions. For situations in which the evidence is inadequate, incomplete, and/or inconsistent, this chapter suggests ways to blend the best available evidence with less formal sources that can bring tacit knowledge and the experience of professionals and other stakeholders to bear.

A large number of individual questions can, of course, be raised by those undertaking efforts to address obesity or other complex public health challenges. Petticrew and Roberts (2003) place such questions into eight broad categories: effectiveness

(Does this work?), process of delivery (How does it work?), salience (Does it matter?), safety (Will it do more good than harm?), acceptability (Will people be willing to use the intervention?), cost-effectiveness (Is it worth buying this service?), appropriateness (Is this the right service/intervention for this group?), and satisfaction (Are stakeholders satisfied with the service?). To this categorization the committee has added such questions as How many and which people are affected? and What is the seriousness of the problem? In Chapter 5, the committee adopts this approach but places these questions in the broad categories of "Why," "What," and "How" and gives a number of examples for each category (Tables 5-1 through 5-3).

Certain types of evidence derived from various study designs could be used to answer some of these questions but not others (Flay et al., 2005). For example, to ascertain the prevalence and severity of a condition and thus the population burden, one needs survey or other surveillance data, not an RCT. To ascertain efficacy, effectiveness, or cost-effectiveness, an RCT may be the best design. To understand how an intervention works, qualitative designs may be the most valuable and appropriate (MacKinnon, 2008). To assess the organizational adoption and practitioner implementation and maintenance of a practice, longitudinal studies of organizational policies and their implementation and enforcement (i.e., studies of quality improvement) may be needed.

As discussed in previous chapters, to assess interventions designed to control obesity at the community level or in real-world settings, RCTs may not be feasible or even possible, and other types of evidence are more appropriate (Mercer et al., 2007; Sanson-Fisher et al., 2007; Swinburn et al., 2005). To apply the terminology adopted for this report (Chapter 5) (Rychetnik et al., 2004), for "Why" (e.g., burden of obesity) or in some cases "How" (e.g., translation of an intervention) questions, RCTs are not the appropriate study design. The same may be true even for some "What" questions (e.g., effectiveness of an intervention) that lend themselves more to formal intervention studies.

Also as discussed in previous chapters, decision makers need to recognize the interrelated nature of factors having an impact on the desired outcome of complex public health interventions. They should view an intervention in the context in which it will be implemented, taking a systems perspective (see Chapter 4). Such a perspective, which evolved from an appreciation of the importance of effectiveness in real-world conditions or natural settings (Flay, 1986), is clearly needed when decision makers evaluate generalizability, as well as level of certainty, in judging the quality of evidence (Green and Glasgow, 2006; Rychetik et al., 2004; Swinburn et al., 2005).

IDENTIFICATION OF APPROPRIATE OUTCOMES

Appropriate outcomes may be multiple and may be short-term, intermediate, or long-term in nature. Regardless, they should be aligned with user needs and interests. For policy makers, for example, the outcomes of interest may be those for which they will be held accountable, which may or may not be directly related to reductions in obesity. In a political context, policy makers may want to know how voters will react, how parents will react, what the costs will be, or whether the ranking of the city or state on body mass index (BMI) levels will change. Health plan directors may want evidence of comparative effectiveness (i.e., comparing the benefits and harms of a competitive intervention in real-world settings) to make decisions on coverage. In any situation with multiple outcomes, which is the usual case, trade-offs may have to be made between these outcomes. For example, an outcome may be cost-effective but not politically popular or feasible. Further discussion of trade-offs can be found later in the chapter.

Logic models are helpful in defining appropriate evaluation outcomes and providing a framework for evaluation. For a long-term outcome, a logic model is useful in defining the short-term and intermediate steps that will lead to that outcome. Outcomes can be goals related to the health of the population (e.g., reduced mortality from diabetes), structural change (e.g., establishment of a new recreation center), a new policy (e.g., access to fresh fruits and vegetables in a Special Supplemental Nutrition Program for Women, Infants, and Children [WIC] program), or others. A recent report by the Institute of Medicine (IOM) (2007) introduces a general logic model for evaluating obesity prevention interventions (for children) (see Chapter 2, Figure 2-2) and applies it specifically to distinct end users, such as government and industry (see Figures 6-2 and 6-3, respectively). This model takes into account the

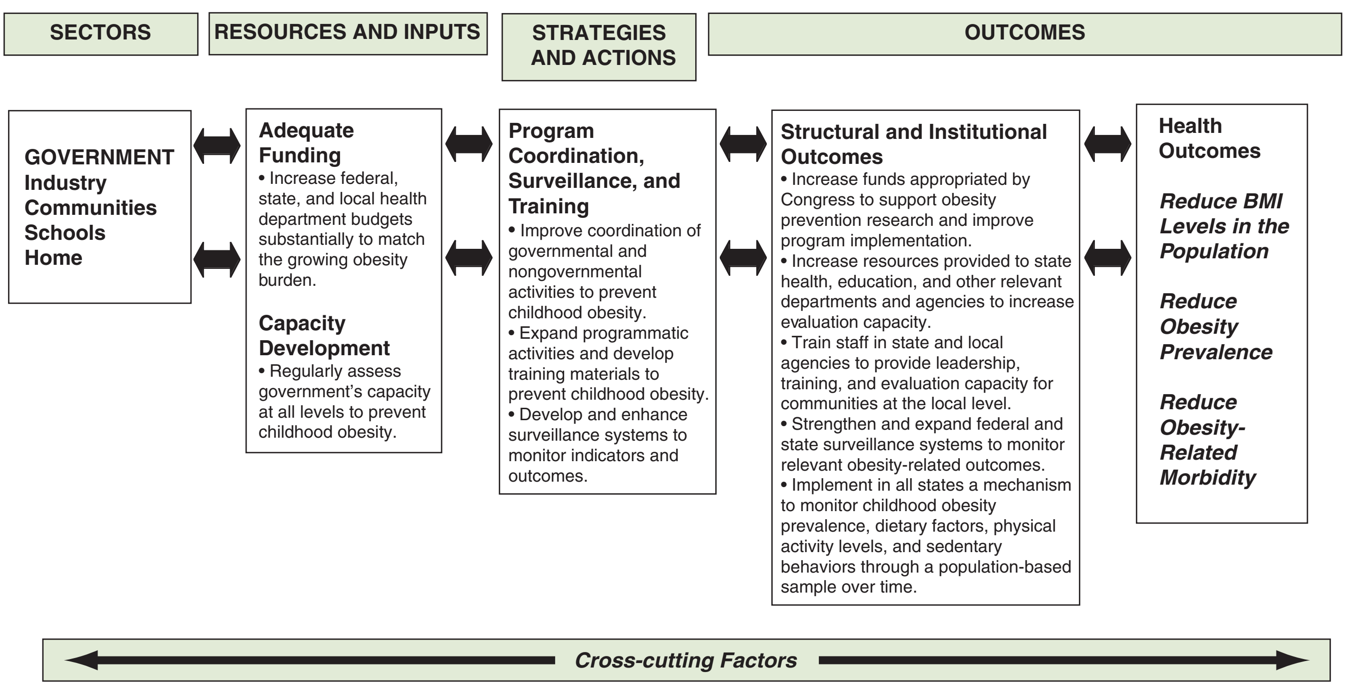


FIGURE 6-2 Evaluation framework for government efforts to support capacity development for preventing childhood obesity.
SOURCE: IOM, 2007.

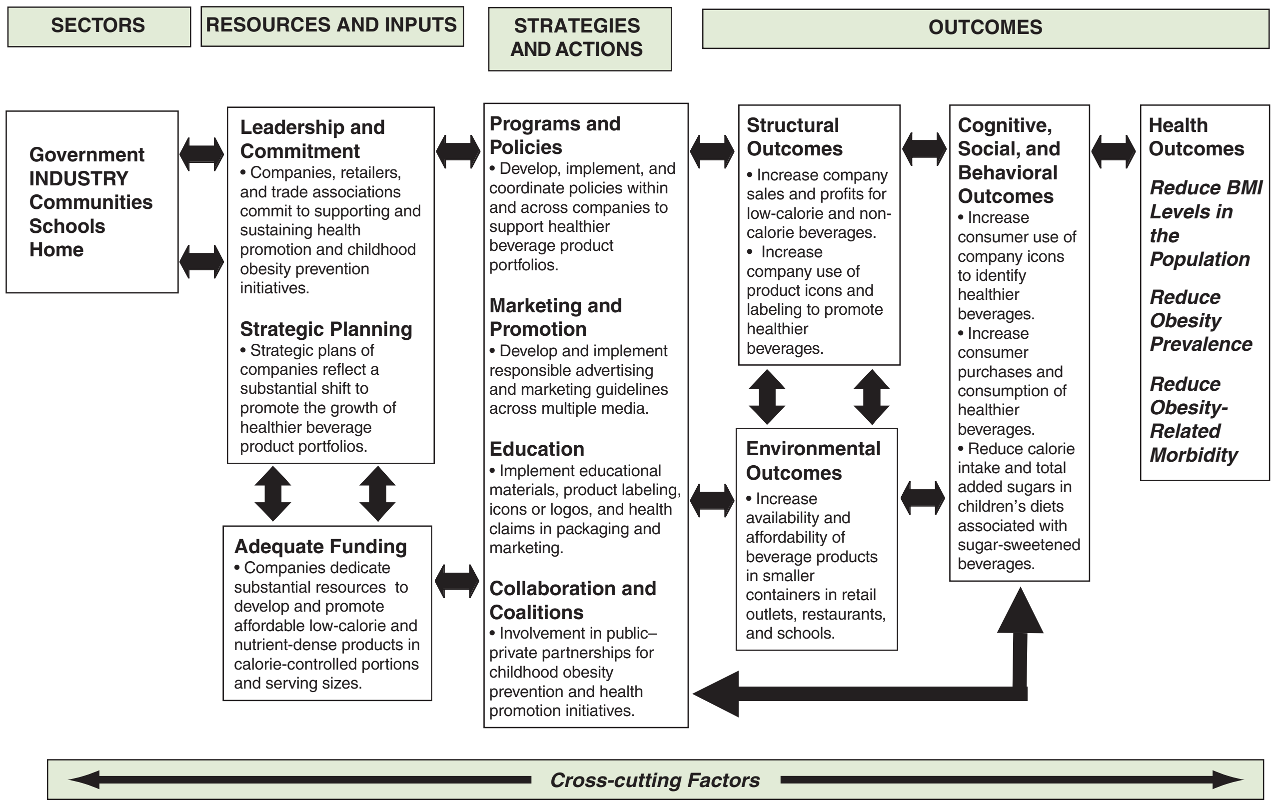


FIGURE 6-3 Evaluation framework for industry efforts to develop low-calorie and nutrient-dense beverages and promote their consumption by children and youth.
SOURCE: IOM, 2007.

interconnected factors that influence the potential impact of an intervention. It facilitates the identification of resources (e.g., funding), strategies and actions (e.g., education, programs), outcomes (e.g., environmental, health), and other cross-cutting factors (e.g., age, culture, psychosocial status) that are important to obesity prevention for particular users.

GENERALIZABILITY AND CONTEXTUAL CONSIDERATIONS

Existing standards of evidence formulate the issue of generalizability in terms of efficacy, effectiveness, and readiness for dissemination (Flay et al., 2005). From this perspective, among the questions to be answered in evaluating whether studies are more or less useful as a source of evidence are the following: How representative were the setting, population, and circumstances in which the studies were conducted? Can the evidence from a study or group of studies be generalized to the multiple settings, populations, and contexts in which the evidence would be applied? Are the interventions studied affordable and scalable in the wide variety of settings where they might be needed, given the resources and personnel available in those settings? For decision makers, the generalizability of evidence is what they might refer to as "relevance": Is the evidence, they ask, relevant to our population and context? Answering this question requires comparing the generalizability of the studies providing the evidence and the context (setting, population, and circumstances) in which the evidence would be applied.

Glasgow and others have called for criteria with which to judge the generalizability of studies in reporting evidence, similar to the Consolidated Standards of Reporting Trials (CONSORT) reporting criteria for RCTs and the Transparent Reporting of Evaluations with Nonrandomized Designs (TREND) quality rating scales for nonrandomized trials (Glasgow et al., 2006a). Box 6-1 details four dimensions of generalizability (using the term "external validity") in the reporting of evidence in most efficacy trials and many effectiveness trials and the specific indicators or questions that warrant consideration in judging the quality of the research (Green and Glasgow, 2006).

EXISTING APPROACHES TO EVALUATING EVIDENCE

The most widely acknowledged approach for evaluating evidence—one that underlies much of what is considered evidence of causation in the health sciences—is the classic nine criteria or "considerations" of Bradford Hill (Hill, 1965): strength of association, consistency, specificity, temporality, biological gradient, plausibility, coherence, experiment, and analogy. All but one of these criteria emphasize the level of causality, largely because the phenomena under study were organisms whose biology was relatively uniform within species, so the generalizability of causal relationships could be assumed with relative certainty.

The rating scheme of the Canadian Task Force on the Periodic Health Examination (Canadian Task Force on the Periodic Health Examination, 1979) was adopted in the late 1980s by the U.S. Preventive Services Task Force (USPSTF) (which systematically reviews evidence for effectiveness and develops recommendations for clinical preventive services) (USPSTF, 1989, 1996). These criteria establish a hierarchy for the quality of studies that places professional judgment and cross-sectional observation at the bottom and RCTs at the top. As described by Green and Glasgow (2006), these criteria also concern themselves almost exclusively with the level of certainty. "The greater weight given to evidence based on multiple studies than a single study was the main . . . [concession] to external validity (or generalizability), . . . [but] even that was justified more on grounds of replicating the results in similar populations and settings than of representing different populations, settings, and circumstances for the interventions and outcomes" (Green and Glasgow, 2006, p. 128). The Cochrane Collaboration has followed this line of evidence evaluation in its systematic reviews, as has the evidence-based medicine movement (Sackett et al., 1996) more generally in its almost exclusive favoring of RCTs (see Chapter 5). As the Cochrane

Box 6-1
Quality Rating Criteria for External Validity

Reach and Representativeness

- Participation: Are there analyses of the participation rate among potential (1) settings, (2) delivery staff, and (3) patients (consumers)?
- Target Audience: Is the intended target audience stated for adoption (at the intended settings, such as worksites or medical offices) and application (at the individual level)?
- Representativeness (settings): Are comparisons made of the similarity of settings in the study to the intended target audience of program settings—or to those settings that decline to participate?
- Representativeness (individuals): Are analyses conducted of the similarity and differences between patients, consumers, or other subjects who participate vs. either those who decline or the intended target audience?

Program or Policy Implementation and Adaptation

- Consistent Implementation: Are data presented on level and quality of implementation of different program components?
- Staff Expertise: Are data presented on the level of training or experience required to deliver the program or quality of implementation by different types of staff?
- Program Adaptation: Is information reported on the extent to which different settings modified or adapted the program to fit their setting?
- Mechanisms: Are data reported on the process(es) or mediating variables through which the program or policy achieved its effects?

Outcomes for Decision Making

- Significance: Are outcomes reported in a way that can be compared to either clinical guidelines or public health goals?
- Adverse Consequences: Do the outcomes reported include quality of life or potential negative outcomes?
- Moderators: Are there any analyses of moderator effects—including of different subgroups of participants and types intervention staff—to assess robustness vs. specificity of effects?
- Sensitivity: Are there any sensitivity analyses to assess dose–response effects, threshold level, or point of diminishing returns on the resources expended?
- Costs: Are data on the costs presented? If so, are standard economic or accounting methods used to fully account for costs?

Maintenance and Institutionalization

- Long-term Effects: Are data reported on longer-term effects, at least 12 months following treatment?
- Institutionalization: Are data reported on the sustainability (or reinvention or evolution) of program implementation at least 12 months after the formal evaluation?
- Attrition: Are data on attrition by condition reported, and are analyses conducted of the representativeness of those who drop-out?

SOURCE: Green, L. W., and R. E. Glasgow, *Evaluation and the Health Professions* 29(1), pp. 126-153, Copyright © 2006 by SAGE Publications. Reprinted by permission of SAGE Publications.

methods have been extended to nonmedical applications, greater acceptability of other types of evidence has been granted, but reluctantly (see below). More recently, the Campbell Collaboration (see Sweet and Moynihan, 2007) attempted to take a related but necessarily distinctive approach to systematic reviews of more complex interventions addressing social problems beyond health, in the arenas of education, crime and justice, and social welfare. The focus was on improving the usefulness of systematic reviews for researchers, policy makers, the media, interest groups, and the broader community of decision makers. The Society for Prevention Research has extended efforts to establish standards for identifying effective prevention programs and policies by issuing standards for efficacy (level of certainty), effectiveness (generalizability), and dissemination (Flay et al., 2005).

The criteria of the USPSTF mentioned above were adapted by the Community Preventive Services Task Force, with greater concern for generalizability in recognition of the more varied public health circumstances of practice beyond clinical settings (Briss et al., 2000, 2004; Green and Kreuter, 2000). The Community Preventive Services Task Force, which is overseeing systematic reviews of interventions designed to promote population health, is giving increasing attention to generalizability in a standardized section on "applicability." Numerous textbooks on research quality have tended to concern themselves primarily with designs for efficacy rather than effectiveness studies, although the growing field of evaluation has increasingly focused on issues of practice-based, real-time, ordinary settings (Glasgow et al., 2006b; Green and Lewis, 1986, 1987; Green et al., 1980). Finally, in the field of epidemiology, Rothman and Greenland (2005) offer a widely cited model that describes causality in terms of sufficient causes and their component causes. This model illuminates important principles such as multicausality, the dependence of the strength of component causes on the prevalence of other component causes, and the interactions among component causes.

The foregoing rules or frameworks for evaluating evidence have increasingly been taken up by the social service professions, building not just on biomedical traditions but also on agricultural and educational research in which experimentation predated much of the action research in the social and behavioral sciences. The social service and education fields have increasingly utilized RCTs, but have faced growing resistance to their limitations and the "simplistic distinction between strong and weak evidence [that] hinged on the use of randomized controlled trials . . ." (Chatterji, 2007, p. 239; see also Hawkins et al., 2007; Mercer et al., 2007; Sanson-Fisher et al., 2007), especially when applied to complex community interventions.

Campbell and Stanley's (1963) widely used set of "threats to internal validity (level of certainty)" for experimental and quasi-experimental designs were accompanied by their seldom referenced "threats to external validity (generalizability)." "The focus on internal validity (level of certainty) was justified on the grounds that without internal validity, external validity or generalizability would be irrelevant or misleading, if not impossible" (Green and Glasgow, 2006, p. 128). These and other issues concerning the level of certainty and generalizability are discussed in greater detail in Chapter 8.

A PROPOSED APPROACH TO EVALUATING THE QUALITY OF SCIENTIFIC EVIDENCE

Scientists have always used criteria or guidelines to organize their thinking about the nature of evidence. Much of what we think we know about the causes of obesity and the current obesity epidemic, for example, is based on the evaluation of evidence using existing criteria. In thinking about the development of a contemporary framework to guide decision making in the complex settings of public health, however, the committee decided to advance a broader view of appropriate evaluation criteria. As described in 2005 in a seminal report from the Institute of Medicine (IOM), these decisions need to be made with the "best available evidence" and cannot wait for the "best possible evidence" or all the desirable evidence to be at hand (IOM, 2005, p. 3). The L.E.A.D. framework should serve the needs of decision makers focused on the obesity epidemic, but can also provide guidance for those making decisions about complex, multi-factorial public health challenges more generally.

The starting point for explaining the committee's approach to evaluating the quality of evidence for obesity prevention is the seven categories of study designs and different sources of evidence presented in Chapter 5. In Table 6-1, this typology is linked to criteria for judging the quality of evidence, drawing on the concept of "critical appraisal criteria" of Rychetnik and colleagues (Rychetnik et al., 2002, 2004). Generally speaking, different types of evidence from different types of study designs are evaluated by different criteria, all of which can be found in the literature on evaluating the quality of each type of evidence. In all cases, high-quality evidence avoids bias, confounding, measurement error, and other threats to validity whenever possible; however, other aspects of quality come into play within the broader scope of evidence advanced by the L.E.A.D. framework.

Users of the L.E.A.D. framework can refer to any of the various criteria for high-quality evidence depending on the source of evidence they have located, following the guidance provided in Chapter 5 as well as the references cited in Table 6-1. This process requires some time and effort by an individual or multidisciplinary group with some expertise in evaluating evidence. Despite the availability of the criteria listed in Table 6-1, making judgments about the quality of evidence can still be challenging. One recommended approach is the eight-step process advanced by Liddle and colleagues (1996):

1. "Select reviewers(s) and agree on details of the review procedure.
2. Specify the objective of the review of evidence.
3. Identify strategies to locate the full range of evidence including unpublished results and work in progress.
4. Classify the literature according to general purpose and study type.

TABLE 6-1 A Typology of Study Designs and Quality Criteria

<table>
<tr><th>Sources of Evidence (research designs, tools, and methods for evidence gathering)</th><th>Existing Criteria for Assessing Quality of Evidence</th></tr>
<tr><td>Nonexperimental or Observational Studies</td><td>Can be assessed by criteria grouped by Liddle et al. (1996):
• Descriptive information about the review or study (e.g., type of intervention)
• Study design, implementation, and analysis
• Overall assessment of the study
Transparent Reporting of Evaluations with Nonrandomized Designs (TREND) is a new attempt to establish criteria for nonrandomized intervention studies that has produced a preliminary statement of criteria for judging such studies (Des Jarlais et al., 2004).</td></tr>
<tr><td>Experimental and Quasi-experimental Studies</td><td>Can be graded by assessment of study design, selection bias, confounding; blinding; data collection and classification of outcomes, follow-up, withdrawal and drop-out, and analysis (Rychetnik et al., 2002, as outlined by the Oxford-based Public Health Resource Unit).
Quality of an RCT is based on (Higgins and Green, 2009):
• Assignment to treatment and control groups and blinding
• Degree of potential confounding
• Classification of outcomes and follow-up
• Appropriate analysis (e.g., “intention to treat”)
Study design is evaluated by levels of evidence (as in those of the Canadian Task Force on the Periodic Health Examination or the Task Force on Community Preventive Services and the U.S. Preventive Services Task Force [USPSTF]). Criteria for the USPSTF are summarized by Harris et al. (2001) and updated by Pettiti et al. (2009).</td></tr>
<tr><td>Qualitative Research</td><td>Standardized quality criteria have not been agreed upon, but should reflect the distinctive goals of the research. As an example of criteria, quality may be determined by the audit trail of processes and decisions made and the credibility of the study methods (Rychetnik et al., 2002, Table 3):
• Clarity of objectives and research questions
• Appropriate selection of method to meet aims
• Clear rationale for sampling strategy
• Appropriate use of triangulation
• Audit trail in data collection and analysis
• Explicit research position and role
• Clear basis for findings
• Transferability of findings
• Relevance, usefulness, importance of findings</td></tr>
<tr><td>Mixed-Method Experimental Studies</td><td>Quality criteria for mixed-method research derive from the quality criteria used for quantitative and qualitative designs separately. A 15-point checklist of criteria for mixed-method research and mixed studies reviews is presented by Pluye et al. (2009). Three points on which mixed-method research can be judged are:
• Justification of the mixed-method design
• Combination of qualitative and quantitative data collection–analysis techniques or procedures
• Integration of qualitative and quantitative data or results</td></tr>
</table>

continued

TABLE 6-1 Continued

Sources of Evidence (research designs, tools, and methods for evidence gathering)	Existing Criteria for Assessing Quality of Evidence
Evidence Synthesis Methods	Questions to consider when appraising a systematic review include (Public Health Resource Unit, 2006): • Did the review address a clearly focused question? • Did the review include the right type of study? • Did the reviewers try to identify all relevant studies? • Did the reviewers assess the quality of all the studies included? • If the results of the study were combined, was it reasonable to do so? • How are the results presented, and what are the main results? • How precise are these results? • Can the results be applied to the local population? • Were all important outcomes considered? • Should practice or policy change as a result of the evidence contained in this review?
Parallel Evidence	Quality is determined by the underlying study designs of the parallel evidence sources in the same way that it is determined for the primary evidence.
Expert Knowledge	Questions to consider when appraising expert knowledge include (World Cancer Research Fund and American Institute for Cancer Research, 2007): • Were methods of review and development of recommendations described? • Was expert knowledge (1) derived from an expert panel, (2) derived from an original review of the literature, and (3) based on published peer-reviewed literature specified in a bibliography? A description of the computer-based Delphi Method for utilizing expert knowledge reliably is provided by Turoff and Hiltz (1996). A description of procedures used to quantify expert opinion (using specialized software) is in Garthwaite et al. (2008).

5. Retrieve the full version of evidence available.
6. Assess the quality of the evidence.
7. Quantify the strength of the evidence.
8. Express the evidence in a standard way." (pp. 6-7).

Step 6 includes checklists for assessing the quality of studies depending on their design and purpose (Liddle et al., 1996).

Most biomedical researchers are familiar with the quality criteria that have been used for experimental and observational epidemiological research, but less so with those used for qualitative studies. Quality is not addressed for qualitative research in the checklists offered by Liddle and colleagues (1996), but can be assessed using the same broad concepts of validity and relevance used for quantitative research. However, these concepts need to be applied differently to account for the distinctive goals of such research, so defining a single method for evaluation is not suggested (Cohen and Crabtree, 2008; Patton, 1999). Mays and Pope (2000) summarize "relativist" criteria for quality, similar to the criteria of Rychetnik and colleagues (2002) (see Table 6-1), that are common to both qualitative and quantitative studies. Others have since reported on criteria that can be used to assess qualitative research (Cohen and Crabtree, 2008; Popay et al., 1998; Reis et al., 2007). In addition, guidance on the description and implementation of qualitative (and mixed-method) research, along with a checklist, has been provided by the National Institutes of Health (Office of Behavioral and Social Sciences Research, 2000).

Criteria also exist for evaluating the quality of systematic reviews themselves, whether they are of quantitative or qualitative studies (Goldsmith et al., 2007). In addition to the criteria of the Public Health Resource Unit (2006) listed in Table 6-1, a detailed set of criteria has been compiled by the Milbank Memorial Foundation and the Centers for Disease Control and Prevention (CDC) (Sweet and Moyniham, 2007).

As noted earlier, expert knowledge is frequently considered to be at the bottom of traditional hierarchies that focus on level of certainty, such as that used by the USPSTF. However, expert knowledge can be of value in evaluating evidence and can also be viewed with certain quality criteria in mind (Garthwaite et al., 2008; Harris et al., 2001; Petitti et al., 2009; Turoff and Hiltz, 1996; World Cancer Research Fund and American Institute for Cancer Research, 2007). The Delphi Method was developed to utilize expert knowledge in a reliable and creative way that is suitable for decision making and has been found to be effective in social policy and public health decision making (Linstone and Turoff, 1975); it is a "structured process for collecting and distilling knowledge from a group of experts" through questionnaires interspersed with controlled feedback (Adler and Ziglio, 1996, p. 3). If these quality criteria are taken into account and conflicts of interest are identified and minimized, decision making can benefit substantially from the considered opinion of experts in a particular field or of practitioners, stakeholders, and policy makers capable of making informed judgments on implementation issues (e.g., doctors, lawyers, scientists, or academics able to interpret the scientific literature or specialized forms of data).

Finally, in addition to the main sources of evidence included in Table 6-1, other sources may be of value in decision making. Many are not independent sources, but closer to a surveillance mechanism or a tool for dissemination of evidence. They include simulation models, health impact assessments, program or policy evaluations, policy scans, and legal opinions. For instance, health impact assessments (described in more detail in Chapter 5, under "What" questions) formally examine the potential health effects of a proposed intervention (Cole and Fielding, 2007). An example is Health Forecasting (University of California–Los Angeles School of Public Health, 2009), which uses a web-based simulation model that allows users to view evidence-based descriptions of populations and subpopulations (disparities) to assess the potential effects of policies and practices on future health outcomes. Another such source, policy evaluations, allows studies of various aspects of a problem to be driven by a clear conceptual model. An example is the International Tobacco Control Policy Evaluation Project, a multidisciplinary, multisite, international endeavor that aims to evaluate and understand the impact of tobacco control policies as they are implemented in countries around the world (Fong et al., 2006). These sources may provide evidence for which there are quality criteria to consider, but are not addressed in detail here.

WHEN SCIENTIFIC EVIDENCE IS NOT A PERFECT FIT: TRADE-OFFS TO CONSIDER

Trade-offs may be involved in considering the quality of various types of evidence available to answer questions about complex, multilevel public health interventions (Mercer et al., 2007). Randomization at the individual level and experimental controls may remain the gold standard, but as pointed out above, these methods are not always possible in population health settings, and they are sometimes counterproductive with respect to the artificial conditions used to implement randomization and control procedures. Therefore, some of the advantages of RCTs may have to be traded off to obtain the best available evidence for decision making. Because no one study is usually sufficient to support decisions on public health interventions, the use of multiple types of evidence (all of good quality for their design) may be the best approach (Mercer et al., 2007), a point further elaborated upon in Chapter 8.

REFERENCES

Adler, M., and E. Ziglio. 1996. *Gazing into the oracle: The Delphi method and its application to social policy and public health*. London, UK: Jessica Kingsley.

Briss, P. A., S. Zaza, M. Pappaioanou, J. Fielding, L. Wright-De Aguero, B. I. Truman, D. P. Hopkins, P. D. Mullen, R. S. Thompson, S. H. Woolf, V. G. Carande-Kulis, L. Anderson, A. R. Hinman, D. V. McQueen, S. M. Teutsch, and J. R. Harris. 2000. Developing an evidence-based Guide to Community Preventive Services—methods. *American Journal of Preventive Medicine* 18(1, Supplement 1):35-43.

Briss, P. A., R. C. Brownson, J. E. Fielding, and S. Zaza. 2004. Developing and using the Guide to Community Preventive Services: Lessons learned about evidence-based public health. *Annual Review of Public Health* 25:281-302.

Campbell, D. T., and J. C. Stanley. 1963. *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally.

Canadian Task Force on the Periodic Health Examination. 1979. The periodic health examination. *Canadian Medical Association Journal* 121(9):1193-1254.

Chatterji, M. 2007. Grades of evidence: Variability in quality of findings in effectiveness studies of complex field interventions. *American Journal of Evaluation* 28(3):239-255.

Cohen, D. J., and B. F. Crabtree. 2008. Evaluative criteria for qualitative research in health care: Controversies and recommendations. *Annals of Family Medicine* 6(4):331-339.

Cole, B. L., and J. E. Fielding. 2007. Health impact assessment: A tool to help policy makers understand health beyond health care. *Annual Review of Public Health* 28:393-412.

Des Jarlais, D. C., C. Lyles, and N. Crepaz. 2004. Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: The TREND statement. *American Journal of Public Health* 94(3):361-366.

Flay, B. R. 1986. Efficacy and effectiveness trials (and other phases of research) in the development of health promotion programs. *Preventive Medicine* 15(5):451-474.

Flay, B. R., A. Biglan, R. F. Boruch, F. G. Castro, D. Gottfredson, S. Kellam, E. K. Moscicki, S. Schinke, J. C. Valentine, and P. Ji. 2005. Standards of evidence: Criteria for efficacy, effectiveness and dissemination. *Prevention Science* 6(3):151-175.

Fong, G. T., K. M. Cummings, R. Borland, G. Hastings, A. Hyland, G. A. Giovino, D. Hammond, and M. E. Thompson. 2006. The conceptual framework of the International Tobacco Control (ITC) Policy Evaluation Project. *Tobacco Control* 15(Supplement 3): iii1-iii2.

Garthwaite, P. H., J. B. Chilcott, D. J. Jenkinson, and P. Tappenden. 2008. Use of expert knowledge in evaluating costs and benefits of alternative service provisions: A case study. *International Journal of Technology Assessment in Health Care* 24(3):350-357.

Glasgow, R., L. Green, L. Klesges, D. Abrams, E. Fisher, M. Goldstein, L. Hayman, J. Ockene, and C. Orleans. 2006a. External validity: We need to do more. *Annals of Behavioral Medicine* 31(2):105-108.

Glasgow, R. E., L. M. Klesges, D. A. Dzewaltowski, P. A. Estabrooks, and T. M. Vogt. 2006b. Evaluating the impact of health promotion programs: Using the RE-AIM framework to form summary measures for decision making involving complex issues. *Health Education Research* 21(5):688-694.

Goldsmith, M. R., C. R. Bankhead, and J. Austoker. 2007. Synthesising quantitative and qualitative research in evidence-based patient information. *Journal of Epidemiology and Community Health* 61(3):262-270.

Green, L. W., and R. E. Glasgow. 2006. Evaluating the relevance, generalization, and applicability of research: Issues in external validation and translation methodology. *Evaluation & the Health Professions* 29(1):126-153.

Green, L. W., and M. W. Kreuter. 2000. Commentary on the emerging Guide to Community Preventive Services from a health promotion perspective. *American Journal of Preventive Medicine* 18(1 Supplement 1):7-9.

Green, L. W., and M. W. Kreuter. 2005. *Health program planning: An educational and ecological approach*. 4th ed. New York: McGraw-Hill.

Green, L. W., and F. M. Lewis. 1986. *Measurement and evaluation in health education and health promotion*. Palo Alto, CA: Mayfield Publishing Company.

Green, L., and F. M. Lewis. 1987. Data analysis in evaluation of health education: Towards standardization of procedures and terminology. *Health Education Research* 2(3):215-221.

Green, L. W., F. M. Lewis, and D. M. Levine. 1980. Balancing statistical data and clinician judgments in the diagnosis of patient educational needs. *Journal of Community Health* 6(2):79-91.

Harris, R. P., M. Helfand, S. H. Woolf, K. N. Lohr, C. D. Mulrow, S. M. Teutsch, and D. Atkins. 2001. Current methods of the U.S. Preventive Services Task Force: A review of the process. *American Journal of Preventive Medicine* 20(3, Supplement):21-35.

Hawkins, N. G., R. W. Sanson-Fisher, A. Shakeshaft, C. D'Este, and L. W. Green. 2007. The multiple baseline design for evaluating population-based research. *American Journal of Preventive Medicine* 33(2):162-168.

Higgins, J. P. T., and S. Green (editors). 2009. *Cochrane handbook for systematic review of interventions,* Version 5.0.2 [updated September 2009]. The Cochrane Collaboration, 2008. http://www.cochrane-handbook.org (accessed December 13, 2009).

Hill, A. B. 1965. The environment and disease: Association or causation. *Proceedings of the Royal Society of Medicine* 58:295-300.

IOM (Institute of Medicine). 2005. *Preventing childhood obesity: Health in the balance.* Edited by J. Koplan, C. T. Liverman, and V. I. Kraak. Washington, DC: The National Academies Press.

IOM. 2007. *Progress in preventing childhood obesity: How do we measure up?* Edited by J. Koplan, C. T. Liverman, V. I. Kraak, and S. L. Wisham. Washington, DC: The National Academies Press.

Liddle, J., M. Williamson, and L. Irwig. 1996. *Method for evaluating research and guideline evidence.* Sydney: NSW Health Department.

Linstone, H. L., and M. Turoff. 1975. *The Delphi method: Techniques and applications.* Reading, MA: Addison-Wesley.

MacKinnon, D. P. 2008. *An introduction to statistical meditation analysis.* New York: Lawrence Erlbaum Associates.

Mays, N., and C. Pope. 2000. Qualitative research in health care: Assessing quality in qualitative research. *British Medical Journal* 320(7226):50-52.

Mercer, S. L., B. J. DeVinney, L. J. Fine, L. W. Green, and D. Dougherty. 2007. Study designs for effectiveness and translation research: Identifying trade-offs. *American Journal of Preventive Medicine* 33(2):139-154.

Office of Behavioral and Social Sciences Research. 2000. Qualitative methods in health research: Opportunities and considerations in application and review. Produced by the NIH Culture and Qualitative Research Interest Group, based on discussions and written comments from the expert working group at a workshop sponsored by the Office of Behavioral and Social Sciences Research. Bethesda, MD: Office of Behavioral and Social Sciences Research.

Patton, M. Q. 1999. Enhancing the quality and credibility of qualitative analysis. *Health Services Research* 34(5, Part 2):1189-1208.

Petitti, D. B., S. M. Teutsch, M. B. Barton, G. F. Sawaya, J. K. Ockene, and T. Dewitt. 2009. Update on the methods of the U.S. Preventive Services Task Force: Insufficient evidence. *Annals of Internal Medicine* 150(3):199-205.

Petticrew, M., and H. Roberts. 2003. Evidence, hierarchies, and typologies: Horses for courses. *Journal of Epidemiology and Community Health* 57(7):527-529.

Pluye, P., M. P. Gagnon, F. Griffiths, and J. Johnson-Lafleur. 2009. A scoring system for appraising mixed methods research, and concomitantly appraising qualitative, quantitative and mixed methods primary studies in mixed studies reviews. *International Journal of Nursing Studies* 46(4):529-546.

Popay, J., A. Rogers, and G. Williams. 1998. Rationale and standards for the systematic review of qualitative literature in health services research. *Qualitative Health Research* 8(3):341-351.

Public Health Resource Unit. 2006. *Critical appraisal skills programme (CASP). Making sense of evidence.* http://www.phru.nhs.uk/Doc_Links/S.Reviews%20Appraisal%20Tool.pdf (accessed December 17, 2009).

Reis, S., D. Hermoni, R. Van-Raalte, R. Dahan, and J. M. Borkan. 2007. Aggregation of qualitative studies—from theory to practice: Patient priorities and family medicine/general practice evaluations. *Patient Education and Counseling* 65(2):214-222.

Rothman, K. J., and S. Greenland. 2005. Causation and causal inference in epidemiology. *American Journal of Public Health* 95(Supplement 1):S144-S150.

Rychetnik, L., M. Frommer, P. Hawe, and A. Shiell. 2002. Criteria for evaluating evidence on public health interventions. *Journal of Epidemiology and Community Health* 56(2):119-127.

Rychetnik, L., P. Hawe, E. Waters, A. Barratt, and M. Frommer. 2004. A glossary for evidence based public health. *Journal of Epidemiology and Community Health* 58(7):538-545.

Sackett, D. L., and J. E. Wennberg. 1997. Choosing the best research design for each question. *British Medical Journal* 315(7123):1633-1640.

Sackett, D. L., W. M. C. Rosenberg, J. A. M. Gray, R. B. Haynes, and W. S. Richardson. 1996. Evidence based medicine: What it is and what it isn't. *British Medical Journal* 312(7023):71-72.

Sanson-Fisher, R. W., B. Bonevski, L. W. Green, and C. D'Este. 2007. Limitations of the randomized controlled trial in evaluating population-based health interventions. *American Journal of Preventive Medicine* 33(2):155-161.

Sweet, M., and R. Moynihan. 2007. *Improving population health: The uses of systematic reviews.* New York: Milbank Memorial Fund and Centers for Disease Control and Prevention.

Swinburn, B., T. Gill, and S. Kumanyika. 2005. Obesity prevention: A proposed framework for translating evidence into action. *Obesity Reviews* 6(1):23-33.

Turoff, M., and S. R. Hiltz. 1996. Computer-based Delphi process. In *Gazing into the oracle: The Delphi method and its application to social policy and public health*, edited by M. Adler and E. Ziglio. London, UK: Jessica Kingsley. Pp. 56-85.

University of California–Los Angeles School of Public Health. 2009. *Health forecasting.* http://www.health.forcasting.org (accessed November 9, 2009).

USPSTF (U.S. Preventive Services Task Force). 1989. *Guide to clinical preventive services.* Baltimore, MD: Lippincott Williams & Wilkins.

USPSTF. 1996. *Guide to clinical preventive services.* 2nd ed. Baltimore, MD: Williams & Wilkins.

World Cancer Research Fund and American Institute for Cancer Research. 2007. *Food, nutrition, physical activity, and the prevention of cancer: A global perspective.* Washington, DC: American Institute for Cancer Research.

7

Assembling Evidence and Informing Decisions

KEY MESSAGES

- When decision makers are deciding on obesity prevention interventions, they need to understand the state of the available evidence pertaining to that decision. This includes evidence on the problem to be addressed, the likely effectiveness and impact of potential interventions, and key considerations involved in their implementation.
- In the face of an emergent, and urgent, public health problem such as obesity, decisions to act often must be made in the face of a relative absence of evidence, or with evidence that is inconclusive, inconsistent, or incomplete.
- If obesity prevention actions must be taken when evidence is limited, this incomplete evidence can be blended carefully and transparently with information from theory, expert opinion, experience, and local wisdom to make the best possible decision.
- Successful evidence gathering, evaluation, and synthesis for use in obesity prevention usually involves a number of disciplines, each with its own methodologies and technical language. The development of a uniform language and structure for summarizing and communicating the relevant evidence in a systematic, transparent, and cross-disciplinary way is critical.

Once the best available evidence has been located (Chapter 5) and evaluated (Chapter 6), it should be assembled and communicated so it can be used to inform decisions (see Figure 7-1). The goal of such evidence reports is to present past experiences in useful, relevant, and readily understood terms to help decision makers choose which programs or policies to implement (i.e., an intervention or a combination of a set of interventions). Within their own disciplines, researchers, program

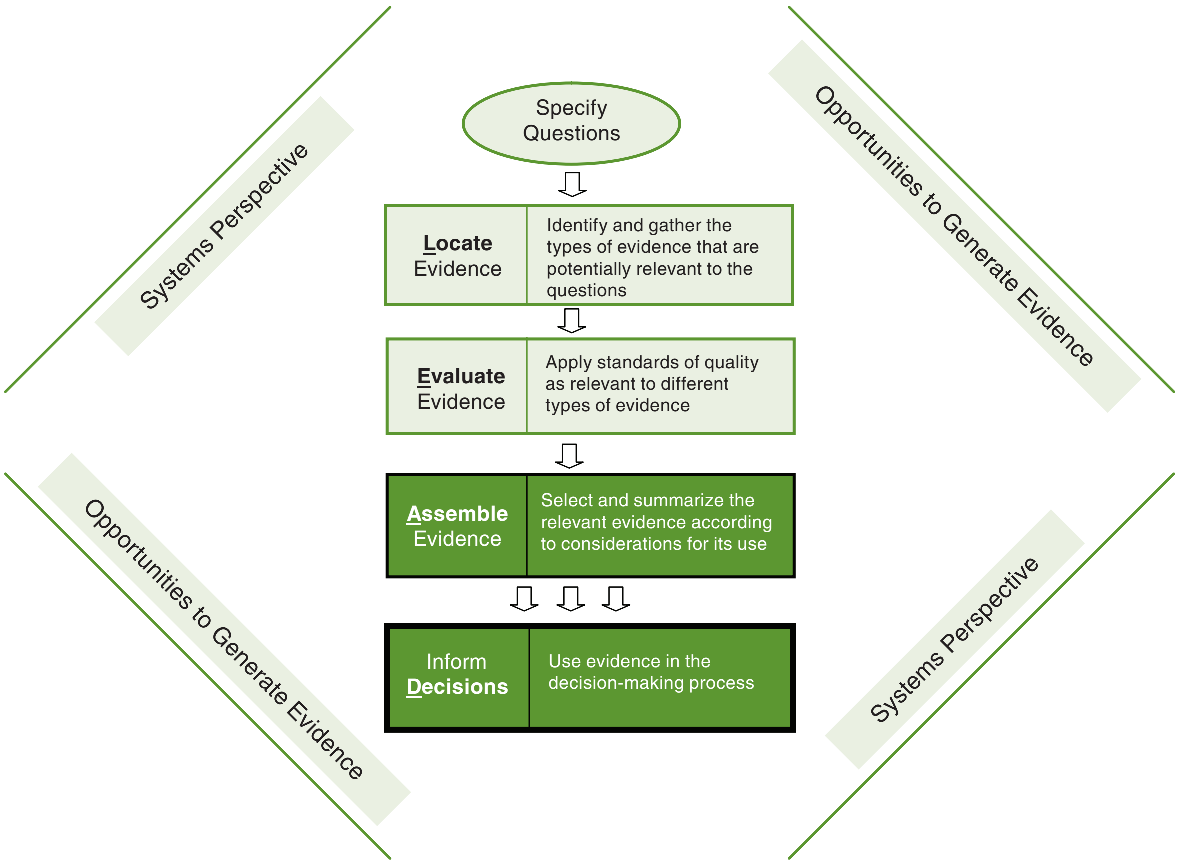


FIGURE 7-1 The **L**ocate Evidence, **E**valuate Evidence, **A**ssemble Evidence, Inform **D**ecisions (L.E.A.D.) framework for obesity prevention decision making.
NOTE: The elements of the framework addressed in this chapter are highlighted.

evaluators, public health officials, and policy makers tend to communicate in highly specialized, useful shorthand. Reporting evidence in accordance with the L.E.A.D. framework will create a shared understanding of the state of relevant knowledge that will support cross-disciplinary decision making. It is also important to recognize that decision makers grappling with an emergent problem such as obesity may need to decide on a course of action in the relative absence of evidence or on the basis of inconclusive, inconsistent, or incomplete evidence. In some cases, the conclusions that can be drawn from the existing evidence may not appear to fit the population, setting, or circumstances at hand. In this context, transparent and structured processes for assembling and reporting evidence allow for reasoned discussion and for modification as new information comes to light. Finally, once a course of action has been chosen, the decision itself and the knowledge gained in the process need to be incorporated into the context of the organization or system where the decision was made. Such "knowledge integration" is the desired final outcome of the application of the L.E.A.D. framework.

The committee suggests that, to inform decisions using the framework proposed in this volume, the conclusions reported to decision makers be formulated to respond to the three questions shown in Figure 5-2 in Chapter 5: (1) Why should we do something about this problem in our situation? (2) What specifically should we do about this problem? and (3) How do we implement this information for our situation? Translating these questions into the language of evidence-based decision making results in three topics that are addressed in this chapter:

- evidence concerning the characteristics of the problem to be addressed in the population targeted by an intervention ("Why" questions);
- evidence supporting the theoretical underpinnings of each potential intervention, presented as a logic model, and evidence supporting the impact of the intervention ("What" questions); and
- considerations involved in implementation, with an appropriate balance of fidelity and adaptation to achieve optimum results ("How" questions).

While each of these topics has been considered in previous chapters, two major sections form the balance of this chapter: (1) specific guidance for assembling the evidence and other information (theory, professional experience, and local wisdom) to support decision making and (2) a standard template for summarizing the evidence, using the L.E.A.D. framework, to provide information relevant to the needs of decision makers. The third section briefly addresses knowledge integration.

GUIDANCE FOR ASSEMBLING THE EVIDENCE

Why Should We Do Something About This Problem in Our Situation?

An evidence report should begin with a succinct summary of the problem to be addressed in the decision-making context. In the case of local decision making, for example, this problem statement may be developed most readily by conducting a community assessment, as discussed in the Institute of Medicine (IOM) report *Local Government Actions to Prevent Childhood Obesity* (IOM, 2009). Community assessments begin with a detailed description of the local conditions to be addressed by an intervention. Local academic partners, public health departments, academic settings, or health information exchanges may be able to provide access to available data and necessary epidemiological or statistical support. Community assessments can also include examining the access of various segments of the community to an optimal environment to support healthy weight, including healthful food and opportunities for physical activity.

The systems thinking discussed in Chapter 4 is important in addressing "why" something should be done about a problem in a particular context. Changes in zoning codes or in agricultural policies, for example, have far-reaching consequences that are best understood by thinking of the larger system implications, often using complex systems models. The information gathered to inform decision makers in those cases would necessarily include both empirical data and models that identify intended effects, dependencies, and potential adverse effects. For these reasons, answers to "Why" questions may require a critique of the model used or the underlying context.

What Specifically Should We Do About This Problem?

Estimates of the effectiveness or impact of an intervention within a well-defined population or setting are essential for assessing its potential to address a public health problem such as obesity. To this end, the evidence should be synthesized for each potential intervention to answer three questions. First, what is the broader context for the intervention? This question should be addressed using a systems perspective. Second, what does the evidence say about the effects of the intervention? Finally, what will be the overall public health impact of the intervention? Decision makers may need to weigh highly effective interventions that reach only a modest number of people against less effective interventions that reach the entire population at risk; they may decide to bring promising interventions to scale even in the face of incomplete information.

The Systems Perspective: Logic Models and the Complexity of Interventions

Reducing the proportion of the population that is overweight or obese is the ultimate goal of all obesity prevention activities; reducing individual caloric intake and/or increasing individual physical activity among at-risk persons are intermediate outcomes required to achieve this goal. Direct measurement of obesity reduction may be infeasible, or may require the passage of time or the simultaneous adoption of multiple policy and program innovations. As a result, most research studies and program evaluations focus on the intermediate outcomes, based on their logical link to the prevention of obesity.

Program evaluations and other evidence synthesis methods (e.g., a mixed-method realist review, discussed below) typically require evaluation designs that incorporate testable theories, technically referred to as logic models. The simplest logic model involves the direct chain of causation of an individual intervention. Predicting the effect of a population-level behavioral intervention on the incidence of obesity should take into account three causal links. First, does the intervention affect behavior? Second, does the change in behavior—presumably a reduction of caloric intake or an increase in physical activity—lead to the desired outcome? Third, does the change in behavior lead to other outcomes, and are they adverse or favorable? At best, the causal links that define each of these relationships are imperfect, which means that predicting the effect of an intervention will always be an inexact science.

Multiple interventions often are required to achieve improvements in population health, and studying the effects of individual interventions will be inadequate to determine their overall effectiveness. As discussed in Chapter 4, the Foresight Group performed an analysis of the many social, economic, and political changes that have contributed to both increasing caloric intake and declining physical activity in the United Kingdom (UK Government Foresight Programme, 2007). Such analyses support the need to take a systems view to understand the implications of interventions for the outcomes of interest. For the purposes of this chapter, it should be noted that the use of qualitative or mathematical models may help identify key mediating variables, providing decision makers with guidance for the development of a portfolio of interventions that may contribute to improvements in population health.

By way of illustration, even the consideration of what appear to be relatively simple interventions, such as those that involve the provision of clear information to the public, can benefit from using a logic model and a systems perspective. Examples include the Arkansas experience in reporting children's body mass index (BMI) to their parents (Raczynski et al., 2009) and the reporting of calorie content on restaurant menus in New York City (Farley et al., 2009). These interventions matter only if certain other conditions are present: If Arkansas parents are told that their children are overweight, will they make adjustments? How does the public sector ensure that recreational resources are available so parents can encourage their children to engage in more physical activity? In New York, will overweight people change their ordering patterns when they know the calorie content of the food offered in restaurants? Will this public reaction lead fast food chains to decrease the calorie density of the items on their menus? In such cases, the intervention itself is built on a systems model that predicts the answers to these questions and helps decision makers develop coordinated interventions that build on each other. Evidence of effectiveness requires the identification of both immediate and long-term outcomes, with a logic model being used to identify the relationships between them.

Several caveats need to be considered in the use of models to predict the effects of an intervention. Models necessarily mix observations (often limited) with theoretical predictions. The more distal an anticipated outcome, the less certainty may be ascribed to that estimate. An evidence report should include a summary of insights gained by taking a systems perspective in relation to the intervention. Systems summaries will seldom provide precise estimates of effects of a particular intervention or interventions; instead, they may highlight areas of concern or provide decision makers with options for monitoring and evaluating new interventions.

Evaluating the Effectiveness of Interventions

Traditional evidence syntheses emphasize assessment of the effectiveness of potential interventions. Similarly, the L.E.A.D. framework includes a synthesis of the evidence gathered, focused on helping the decision maker understand the likely outcomes of each potential intervention and the degree of certainty with which that projection can be made: What is the likely magnitude of the change to be expected with this intervention alone, or how dependent might the desired change be on combining this intervention with other components in a more comprehensive strategy or policy? How tightly linked are the immediate effects of the intervention to the goals of decreased caloric intake and increased energy expenditure? In addition, careful attention should be paid to the types of information included, as well as the standards used to evaluate each source of evidence, to determine the level of certainty of the causal link between an intervention and the observed outcomes (see Chapters 5 and 6 for further detail).

The relative importance of different types of evidence will depend on the decision-making context. Best and colleagues (2008a) view evidence according to two distinct spectra: an ecological perspective at the individual, organizational, and systems levels and a programmatic perspective within basic/discovery, clinical, and population research settings. The resulting matrix of evidence recognizes both the research setting, or context, in which the evidence was generated and the fact that mechanisms, actors, and actions will differ depending on whether individuals, organizations, or systems are targeted.

In general, the best evidence at the individual level, particularly in the discovery phase of research, will be derived from experimental or quasi-experimental studies (Best et al., 2008b; Rychetik et al., 2004). These approaches, including the use of statistical meta-analysis techniques, may yield quantitative estimates of the effectiveness of individual-level interventions (see Chapter 5 for more discussion of meta-analysis); however, intent-to-treat analysis of RCTs is the best way to minimize bias in ascertaining treatment effects in individual studies (Montori and Guyatt, 2001). Questions concerning systems and populations may best be answered using nonexperimental sources, such as descriptive research or qualitative analyses (Best et al., 2008b).

Program evaluation techniques evolve with program growth and complexity. In anticipating the types of data that may be available on interventions, Jacobs (2003) notes that while all programs should involve some sort of evaluation, evaluation capacities and interests are not static. For purposes of the present discussion, decision makers may expect that newer, smaller programs are more likely to have lower-quality evidence, while larger, more established programs are likely to generate higher-quality data more directly relevant to health outcomes.

Decision makers typically adapt or modify interventions to suit local conditions. This adaptation may influence the reported effectiveness of an intervention. One method used to transform information obtained from published reports into meaningful generalized guidance is a realist review (a type of mixed-method analysis used to synthesize evidence) (Pawson et al., 2005). As described by Greenhalgh and colleagues (2007, p. 858) in a recent Cochrane review of school feeding programs, "[a] realist review exposes and articulates the mechanisms by which the primary studies assumed the interventions to work (either explicitly or implicitly); gathers evidence from primary sources about the process of implementing the intervention; and evaluates that evidence so as to judge the integrity with which each theory was actually tested and (where relevant) adjudicate between different theories." The realist approach allows decision makers to apply prior evidence, even in circumstances where complete fidelity to the original studied intervention (i.e., conducted as planned) may not be practical or possible. In the Greenhalgh and colleagues review, for example, the investigators identify 10 process factors associated with improved efficacy, as well as a number of technical factors that complicate the measurement and assessment of outcomes. In another case, Connelly and colleagues (2007) use the realist review strategy to assess the importance of compulsory aerobic physical education in schools (see Chapter 5 for further discussion).

The intent of evidence synthesis using the L.E.A.D. framework is to build directly on the experience of conducting systematic reviews. As has been described, however, systematic reviews conducted in accordance with this framework will differ from those that rely on quantitative methods for synthesizing the results of experimental data and will gather and evaluate a broader array of available information. The committee's decision to include evidence derived from many types of studies, qualitative data, and alternative data sources hitherto considered unworthy of inclusion is based on recently published evidence. Examining the inclusion and exclusion of observational studies along with randomized controlled trials (RCTs) in meta-analyses led Shrier and colleagues (2007, p. 1203) to the following conclusions: (1) "including information from observational studies may improve the inference based on only randomized trials"; (2) the estimate of effect is similar for meta-analyses based on observational studies and RCTs; (3) the "advantages of including both . . . could outweigh the disadvantages . . ."; (4) "observational studies should not be excluded a priori"; and (5) RCTs have been implemented largely in interventions with individuals, while community-based interventions most frequently use other research methods. Based on technical considerations alone, RCTs are most efficient and reliable when one is considering short-term, readily measured outcomes among individual subjects. While they may be adapted to larger, community-level interventions, doing so greatly increases their cost and complexity and may decrease their reliability (Sege and De Vos, 2008). Chapter 8 presents further detail on the challenges and trade-offs of randomized experiments.

Synthesis of Information from Disparate Sources

The fundamental limitations of even the best available evidence mean that the L.E.A.D. framework will need to provide for ways to fill the gaps in the evidence carefully and transparently using several sources, including theory, professional experience, and local wisdom (see Box 7-1). This, indeed, was the experience of tobacco control efforts across the United States in the last third of the 20th century as the tobacco epidemic was reversed and halved following its peak in 1966 (see Box 7-2). The combining of the above sources with the science-based evidence in the process of planning an intervention has been characterized as "matching, mapping, pooling, and patching" (Green and Glasgow, 2006; Green and Kreuter, 2005, pp. 197-203). Appendix E offers additional in-depth guidance for using theory, professional experience, and local wisdom to adapt science-based evidence to local settings, populations, and times. Included are ways to *match* the evidence with appropriate intervention targets; to *map*

Box 7-1
Additional Sources of Information

Theory—the basis for generalizing from evidence by matching the source of the evidence to the circumstances in which it would be applied. Social and behavioral theory specifies the conditions under which stimulus A (an intervention) produces response B (a change in attitude, belief, behavior, or organizational or community conditions). It allows for consideration of whether such conditions are salient in generalizing the intervention effects found in previous studies to the circumstances in which the intervention would be applied in the particular population, setting, and time at hand.

Professional experience—pooling of the experience of those who have addressed the problem in populations or circumstances similar to those at hand. The result is sometimes referred as tacit knowledge and sometimes as anecdotal evidence. Systematic procedures can make such pooling more reliable (D'Onofrio, 2001).

Local wisdom—engagement in the planning of interventions of those who have lived with the problem.

Box 7-2
Blending Theory, Expert Opinion, and Local Wisdom: Tobacco Control Efforts

In the latter part of the 20th century, innovators in schools, worksites, restaurants, towns, cities, and states enacted clean air policies and conducted smoking prevention and cessation programs, with varying degrees of evidence supporting their actions (Eriksen, 2005; Mercer et al., 2005); similar patterns can be seen in other public health successes of the period, such as injury control (Martin et al., 2007). Those policies and programs that demonstrated effectiveness and sustainability were deemed worthy of note by other jurisdictions. The most notable state tobacco control programs, for example, were those of California and Massachusetts. Those two states, with the support of increased tobacco taxes, had mandates in the early 1990s to undertake and evaluate more comprehensive programs (IOM, 2007). They mobilized aggressively with mass media and programs in schools, worksites, and communities. Evaluation demonstrated first a doubling, then a tripling, of the annual rate of decline in tobacco consumption in California relative to the other 48 states (Siegel et al., 2000). Massachusetts started later than California but eventually achieved a quadrupling of those average 48-state rates of reduced smoking as its program hit its stride in the mid-1990s with higher taxes and greater per capita program expenditures (CDC, 1996). The other states showed more interest in the experience and evaluation data from these two states than in the several thousand randomized controlled trials (RCTs) of smoking cessation and prevention in the scientific literature. This was the case because they perceived the greater relevance and representativeness of the two states' experience relative to the highly controlled, often artificial circumstances and sampled populations of RCTs.

the adaptations of interventions to the characteristics of the setting, population, and time; to use professional consultations to *pool* the experience of other decision makers in addressing populations and settings similar to those at hand; and to use community involvement to *patch* the gaps in evidence, theory, and professional experience by engaging local residents to offer their interpretation of the proposed intervention and their perception of its fit with their needs.

Evaluating the Impact of Interventions

Estimates of the effectiveness of an intervention within a well-defined population or setting provide important information for policy decisions. It is important to use such estimates to begin to assess the potential of an intervention to address a public health problem such as obesity. Although consensus does not exist on the criteria necessary to confirm that an intervention has produced a significant public health impact, the RE-AIM framework provides a comprehensive model for evaluating the potential impact of a public health intervention. The five dimensions that form the framework's acronym are reach, effectiveness, adoption, implementation, and maintenance. *Reach* refers to the size and characteristics of the population affected by an intervention. Public health interventions may be intended to reach all of part of the general population; may be characterized as universal, selective, or indicated based on the targeted population, which is determined on the basis of risk characteristics (as described in Chapter 2); or may be intended to reach those who can help the intended population (e.g., parents of obese children, policy makers). Typically, an intervention will reach only a portion of the intended population. *Effectiveness* is the impact of the intervention on targeted outcomes and quality of life. *Adoption* refers to the proportion and representativeness of settings and intervention staff that agree to deliver a program. *Implementation* is the extent to which the intervention is implemented as intended. Finally, *maintenance* denotes the extent to which a program is sustained over time within the delivering organization.

Consideration of each of the five dimensions of the RE-AIM framework can help evaluators understand the broad array of issues an effective program needs to address. Glasgow and colleagues (2006) propose several combined impact indices and discuss the strengths and weaknesses of each. They suggest that combining the RE-AIM dimensions, rather than considering each individually, would be most useful for making policy decisions. The metrics involved in calculating the impact indices are beyond the scope of this report but are described by the authors.

Within the three broad categories of intended reach (universal, selective, or indicated populations), each intervention may also be examined to determine uptake—the proportion of the intended population that the intervention will actually reach. Consider, for example, two different universal interventions: tax policies may well reach the entire population, although the effect may be variable depending on individual resources, while public awareness campaigns designed to reach an entire community typically reach only a portion of the community, depending on the outreach methods used. A simple intervention that reaches the entire population may have substantial impact; in the case of tobacco control, cigarette taxes proved to have an impact due in large part to their ability to reach all smokers. The impact of other universal tobacco control interventions was achieved with more complex strategies.

It is important to point out that some of the early policy victories in tobacco control did not have a large impact, but they paved the way for subsequent policy actions that had a much greater impact and thus were critical to the success of those actions. For example, school policies prohibiting smoking on school property by teachers and staff affected small numbers of smokers, but set an example for subsequent clean air legislation, smoke-free worksites, and eventually smoke-free restaurants. The policies also helped denormalize smoking in the eyes of school-age children. Another example is the mounting of mass media campaigns that initially had limited impact on smokers themselves, but helped create an informed electorate that would support ballot initiatives and legislative acts for higher tobacco taxes, more clean air controls, and more constraints on advertising and the sale of tobacco to minors. Understanding the context of an intervention and the preliminary actions that preceded it can be key in determining its effectiveness and potential impact.

How Do We Implement This Information for Our Situation?

When making choices about how to address public health problems such as obesity, decision makers should consider whether the relationships identified in research studies will hold up in their own state, locality, or setting. This is especially so if the data on which best-practice evidence is based come from academic settings, college towns, university-affiliated hospitals or clinics, or artificially simulated settings, as is often the case with published evidence. Decision makers should also consider whether the resources and oversight of interventions in the evidence-producing studies are reproducible in their local settings, and whether the best-practice interventions can be taken to scale with larger numbers of organizations or staff lacking the training and close supervision that typically characterize the protocols of scientific studies. Were the extensive training and close supervision of those conducting the intervention related only to the research functions, or are they essential to successful replication of the intervention's effectiveness in any setting? Was the formally controlled trial artificially complex for research purposes, or is its protocol intrinsically high maintenance? In addition, decision makers need to consider what degree of latitude their policies, program plans, or instructions and guidelines can allow in applying the tested interventions to the variety of subpopulations or individual cases that may be encountered. Although the scientific view of many evidence-based guidelines is that they should be applied faithfully and rigorously as in the protocol tested in efficacy trials or effectiveness studies, that is, applied with "fidelity," deviations from such strict adherence to protocol may be necessary for adaptation to local circumstances. Yet if an intervention then does not work—if the results in the application fall short of those in the scientific trials—practitioners are likely to be blamed for deviating from the protocol when other factors, such as a different culture in the target population, may be responsible.

The reality that most who use evidence to support obesity prevention policy or practice are likely to face is that the target population for an obesity prevention intervention will be more diverse than the populations in the original studies (Cohen et al., 2008). This situation often leads to the fidelity problem because users need to adapt the protocol, but it also raises issues related to generalizability that characterize much research; in particular, the implementation of the intervention is highly controlled, and the study population settings, procedures, and circumstance often are not representative of the typical settings or target populations at large.

The generalizability problem relates in part to a distinction between function and form with respect to each component of an intervention (Hawe et al., 2004). *Functional* components are those regarded as essential to successful implementation wherever and with whomever the intervention might be applied, whereas those components of the intervention that are a matter of *form* usually lend themselves to successful adaptation to different populations, settings, and circumstances. While form should in general follow function, it is necessary to decide which components of an intervention are essential for implementation and which can bear modification. Published studies usually do not offer guidance on these decisions. The research community needs to provide more guidance to implementers who are faced with the need to tailor components of an intervention to individuals or to various population segments (cultures, genders, age groups, etc.), settings, and times.

Decision makers should also consider whether the research process, including recruitment and informed consent procedures, screening, and attrition, led to an unrepresentative pool of subjects on which the studies' conclusions were based. This issue is akin to that of the representativeness of the cases on which the conclusions were based. It stems, however, not from the homogeneity or unrepresentativeness of the original population sampled. Rather, sampling is altered by the experimental conditions; the study protocol; and/or dissatisfaction with questionnaires, blood draws, weighing, or other research procedures—all of which could produce an unrepresentative final sample after the initial representative population was properly sampled.

Finally, decision makers should consider that obesity is unlike many of the health-related problems to which the usual canons of evidence-based medicine apply, especially those problems that have a singular causal agent, a straightforward mechanism of causation, or a relatively consistent result for all who would receive a standardized intervention. Obesity should be seen as belonging in a class of problems that are influenced by a constellation of complex social and political factors, some of which change during the process of solving the problem. Such a problem is likely to be viewed differently by different populations, practitioners, and vendors of services and products that influence it, depending on the perspectives and biases of those with a stake in the problem (Kreuter et al., 2004). Therefore, such problems are less amenable to the usual methods of randomized experiments than most of the more strictly biomedical points of intervention tested for evidence-based medicine. This is precisely why a systems perspective is core to the L.E.A.D. framework (see Chapter 4).

In a medical setting, resource evaluation for a potential intervention may be straightforward and based simply on the marginal cost of administering the intervention within an existing structure. For example, the British Grading of Recommendations Assessment, Development, and Evaluation (GRADE) criteria for developing recommendations in a medical setting include resource considerations. The illustrative example offered in a recent description of GRADE assesses the costs of administering magnesium sulfate to prevent pre-eclampsia and provides separate cost/benefit analyses for high-, medium-, and low-income countries (Guyatt et al., 2008a).

The RE-AIM criteria discussed above shift the focus of evaluation of issues of generalizability and reproducibility by setting explicit standards for the reporting of resources used for the implementation and maintenance of interventions (Jilcott et al., 2007). The authors caution that studies often, either by design or inadvertently, indicate additional implementation resources that may not be available outside of a research setting. Armstrong and colleagues (2008) suggest that public health agencies should report data on the resources necessary for program implementation and sustainability, as well as the effectiveness results of pilot interventions.

Other considerations have been proposed for assessing the available evidence with respect to implementation. A notable example is the "filter criteria" for decision makers delineated by Glasziou and Longbottom (1999) and Swinburn and colleagues (2005).

Existing Tools for Assembling Evidence

In the majority of situations encountered, the available data may differ with regard to details of an intervention or the context and/or population in which the intervention would be implemented. In these circumstances, the evidence review should attempt to bridge this gap. Table 7-1 provides brief descriptions and references for existing tools that can be used to assemble the evidence, taking into account the key considerations described above.

A TEMPLATE FOR SUMMARIZING THE EVIDENCE

Evidence summaries should use a uniform language and structure, for many reasons. A uniform language for drawing and describing conclusions signals that the authors used a consistent set of procedures to evaluate and synthesize the evidence. A uniform language also facilitates communication among different disciplines during policy discussions and decision making. Achieving uniformity may sacrifice nuance; on balance, however, it improves the clarity of communication. Clarity is particularly important in complex public health arenas, such as obesity prevention, that necessarily involve experts from many different disciplines, each with its own jargon and standards for communication. Use of a uniform language and structure to report the process and

TABLE 7-1 Existing Tools for Assembling Evidence

Tool	Purpose	Criteria Specific to Assembling Evidence
Realist reviews	Synthesis of evidence using mixed-method analysis to combine qualitative and quantitative studies	Criteria relate to the effectiveness of interventions
Meta-analysis	Pooling of results of experimental and quasi-experimental designs to determine effect size	Criteria relate to the effectiveness of interventions
Grading of Recommendations Assessment, Development, and Evaluation (GRADE)	Classification of the quality of evidence and strength of recommendations in medical settings	Factors that affect the strength of recommendations (Guyatt et al., 2008b): • Quality of evidence • Uncertainty about the balance between desirable and undesirable effects • Uncertainty or variability in values and preferences • Uncertainty about whether the intervention represents a wise use of resources
Guide to Community Preventive Services: Systematic Reviews and Evidence-Based Recommendations	Process for systematically reviewing evidence and translating it into recommendations	Criteria for translating evidence into recommendations (Briss et al., 2000): • Translate the body of evidence of effectiveness into recommendations • Consider information on evidence other than effectiveness (applicability, other effects, economic evaluations, barriers to implementation) • Identify and summarize research gaps
Reach, Effectiveness, Adoption, Implementation, Maintenance (RE-AIM)	Framework for systematically considering strengths and weaknesses of interventions to guide program planning	Criteria for translating research into practice (Kaiser Permanente Institute for Health Research, 2010): • Reach • Effectiveness • Adoption • Implementation • Maintenance
Health Canada framework for identifying, assessing, and managing health risks	Framework for developing program-specific implementation procedures involved in risk management	Criteria for analyzing potential risk management options (Health Canada, 2000): • How quickly the risk must be addressed • Risk vs. benefits • Expected costs (of implementing the option) • Risk, cost, and benefit ratios (efficiency) • Distribution of risks, costs, and benefits (fairness) • Available resources • Unintended consequences • Residual risks (level of risk that remains after the option is implemented) • Perceptions, concerns, and values of interested and affected parties • Acceptability of the risk, the option, and the residual risk to interested and affected parties • Other criteria used for option analysis in similar situations

continued

TABLE 7-1 Continued

Tool	Purpose	Criteria Specific to Assembling Evidence
Obesity prevention evidence framework	Framework for evidence-based obesity prevention decision making	Criteria for selecting a portfolio of policies, programs, and actions (Swinburn et al., 2005): • Certainty as to effectiveness and potential population impact of potential interventions ("promise table") • Feasibility • Sustainability • Effects on equity • Potential side effects • Acceptability to stakeholders
Matching, mapping, pooling, and patching	Combining of theory, professional experience, and local wisdom with the science-based evidence in the process of planning an intervention	Criteria for identifying program components and interventions (Green and Kreuter, 2005): • Matching the evidence with appropriate intervention targets • Mapping the adaptation of an intervention to the setting, population, and time at hand • Pooling information about prior interventions • Patching together evidence-based practices, theory-based programs, and practice-based experiences

conclusions of evidence synthesis makes the information easier for decision makers to understand as they consider each potential intervention or set of interventions that combine individual actions or strategies (both referred to from this point further as reporting on a potential intervention).

As these reports on many different interventions and combinations of interventions begin to proliferate and become broadly available, they will serve as starting points for others facing similar decisions in similar environments. They also are likely to promote productive discussion and debate among decision makers and others, as well as assist researchers and research funders as they decide which kinds of studies are most needed for obesity prevention decision making.

Elements of the Reporting Template

The L.E.A.D. framework calls for summarizing the conclusions of evidence synthesis for each potential intervention by using a uniform reporting template that carefully follows the main tenets of the framework. It is most likely that these reports will be developed by the decision maker's intermediaries—his/her own staff; staff from a particular government department, such as public health; or a group that specializes in gathering and synthesizing evidence. The report should follow that template presented in Box 7-3. The following subsections briefly describe each element of this template and offer specific examples to illustrate its purpose within the decision-making context starting with a brief example of a L.E.A.D. framework report (Box 7-4).

Box 7-3
Elements of the Reporting Template: Using the L.E.A.D. Framework to Inform Decisions

1. Question asked by the decision maker
2. Strategy for locating evidence
3. Evidence table
4. Summary of evidence
    a. Why should we do something about this problem in our situation?
    Summary of local conditions to be addressed by the proposal
    b. What specifically should we do about this problem?
    (1) *Implications of a systems perspective (i.e., rationale)*
    (2) *Likely effectiveness*
    (3) *Likely reach/impact*
    c. How do we implement this information for our situation?
    Summary of lessons learned from previous implementation in other settings and relevant local considerations

Question Asked by the Decision Maker

As described in Chapter 1, there are a wide variety of decision scenarios related to obesity prevention policies and programs. A decision maker may be faced with a decision about a specific, populationwide policy change with a potential direct impact on physical activity or other weight-related behavior of adults or children. For example, a school board member may ask staff to assist in deciding whether to vote for a measure to send parents reports of their children's BMI levels based on annual weight and height screening. Or a decision maker may need to undertake or retain a specific program aimed at increasing access to food or physical activity options in a community. Or the director of a youth service organization may ask staff to help select a family-oriented nutrition and fitness program for nationwide implementation. Finally, a decision maker may need to select one of a set of potential interventions that would decrease obesity-promoting environmental influences; for instance, an employer may ask staff to help determine which wellness services should be included in the employee benefit package next year.

Identifying the question posed by the decision maker is the essential first step in producing the evidence report. This question informs what approach is taken to selecting, implementing, and evaluating interventions; why an intervention is needed; what should be done; how the intervention can work in the given context; and how changes in well-established policies and practices can be justified.

Box 7-4
Example L.E.A.D. Framework Evidence Report

1. Question Asked by the Decision Maker

   State the problem posed by the decision maker. For example:

   This summary was prepared to inform the School Board [decision maker] concerning the issue of whether soda vending machines should be removed from high schools to contribute to the prevention of obesity among our students.

2. Strategy for Locating Evidence

   Summarize the strategy for locating evidence. For example:

   a. To determine the current childhood obesity situation in our city, we accessed data available from the Department of Health and from the public schools.
   b. We asked Dr. Smart from the Public Health School to perform a search using multiple databases, including MEDLINE, the Cochrane Database of Systematic Reviews, and SCOPUS, to determine the available evidence linking soda machines to childhood obesity. Relevant articles were retrieved according to the L.E.A.D. framework methods.
   c. Community leaders (Superintendent of Schools, School Council Chair, and Commissioner of the Department of Health) were asked to review the evidence table and comment on its accuracy and completeness.

This list may include much more evidence from many sources—these are just examples.

3. Evidence Table

   Develop a table summarizing the characteristics of the evidence that was gathered, including its quality. The evidence should be organized by the "Why," "What," "How" typology of the L.E.A.D. framework and its quality assessed according to the methods described in Chapters 5 and 6. See a brief example of an evidence table in Table 7-2.

4. Summary of Evidence

   Synthesize the evidence according to L.E.A.D. framework considerations and methods described in Chapter 7, and summarize it according to the "Why," "What," "How" typology. For example:

   a. **Why should we do something about this problem in our situation?** Data from our Department of Health indicate that rates of obesity in our high school students are among the highest in the state, and current trends demonstrate the problem is getting worse.

Strategy for Locating Evidence

As described in detail in Chapter 5, the strategy for locating evidence should begin with answering the three questions that frame the discussion in this chapter: (1) Why should we do something about this problem in our situation? (2) What specifically should we do about this problem? and (3) How do we implement this information for our situation? Collecting evidence in accordance with this typology will help expand the perspective on the forms of evidence that are potentially relevant to answering the obesity prevention question asked by the decision maker and on the potential sources from which to gather that evidence. The strategy used to locate evidence for each potential intervention depends on the context of the question being asked, as well as

b. **What specifically should we do about this problem?** Remove soda vending machines from the high schools.
   (1) **Rationale:** Our logic model is simple—removing soda machines from the high schools will result in reducing the caloric intake of our high school students and help decrease the rates of obesity. If overweight students consume one can of sugar-sweetened soda per day, the result is an increase in caloric intake of 150 calories per child per day. Obesity results from an imbalance of caloric intake and energy expenditure; if we can reduce caloric intake, the incidence of obesity is likely to fall. Because students obtain their soda largely through purchases at school, they are unlikely to substitute other sweetened beverages for the soda while at school.
   (2) **Effectiveness:** High-quality evidence demonstrates the association between soda consumption and obesity. The IOM (2007) suggests that the only beverages offered to children during the school day should be water, low-fat and nonfat milk, and 100 percent fruit juice. A recent report from Texas demonstrates that removing certain high-fat snacks from cafeterias reduced student consumption of these foods (Cullen and Watson, 2009). We could not find any direct evidence that removing soda machines will reduce student soda consumption, but a recent study conducted in Texas demonstrated reduced snack food consumption when these items were removed from the school cafeterias. The conclusions of that report are likely to hold for this intervention as well. Based on the data collected, we judge this intervention to be PROMISING.
   (3) **Reach/impact:** This intervention will reach all children who attend our public schools, representing 82 percent of the children in the city aged 12-18. If this change results in a 50 percent drop in sugar-sweetened beverage consumption, we will have reduced the average caloric intake of 82 percent of the population by 5 percent.

c. **How do we implement this information for our situation?** Our contract with the local distributor is up for renewal at the end of this school year. If we provide 90 days' notice of our intention not to renew the contract, we will not be required to pay any penalty. In the neighboring county, the distributor called the local paper to complain that the School Board was creating a "nanny state"; however, there was little organized opposition to the change. It will cost approximately $10,000 per school to repair or replace water fountains in each cafeteria, for a total of $60,000 in initial costs. The district will also lose the revenue from soda sales, which totaled $140,000 in fiscal year 2008. Some parents and students will object to the removal on the grounds that high school students are capable of making their own beverage decisions. Principals and teachers should have information available to help them address these concerns.

The entries under a, b, and c above are merely examples. A summary for decision makers is likely to include many more examples and more extensive explanatory information. The purpose of this brief model is to give the reader a concrete sense of what a L.E.A.D. evidence report would look like. Such a report should be prepared for each potential intervention to be considered by the decision maker.

the intermediary's knowledge of and access to study designs and methodologies in diverse fields; therefore, it is important to describe this strategy clearly in the report. (For a list of selected information sources, see Appendix D.)

Evidence Table

As described in Chapter 6, once the evidence for each potential intervention has been located, it should be evaluated. Each type of evidence should be evaluated according to established criteria for assessing the quality of that type of evidence, including the level of certainty and generalizability. In addition, the relative value of each source of evidence depends on the decision-making context. Table 7-2 offers a template for

TABLE 7-2 Evidence Table Template

Example Decision Maker Question: Should our schools change the contents of vending machines to help prevent obesity among our students?

Issue/Potential Intervention	Source	Type	Quality	Outcomes/Findings
Part A: Why should we do something about this problem in our situation?				
All needs assessment evidence for the particular problem (e.g., health burden, frequency/incidence, social determinants, trends, health disparities, monetary and social costs) should be described in this section. One example:				
Childhood obesity rates in our area (2008 data)	Department of Health Expanded Immunization Registry	Quantitative report using existing database	High	A map of our community showing childhood obesity rates by neighborhood is provided.
Part B: What specifically should we do about this problem?				
All evidence associated with the effectiveness and impact of a particular intervention (e.g., causal pathways, outside influences, sustained effects, unintended consequences) should be described in this section. One example:				
School nutrition policy changes	Study citation	Descriptive, quantitative	High	Policy changes in one state's public schools significantly reduced calorie consumption among its youth during the school day.
Part C: How do we implement this information for our situation?				
All evidence associated with the relevance and implementation of a particular intervention (e.g., generalizabiltiy, sustainability, cost-effectiveness, cost feasibility, strategic planning, implementation policies, potential challenges) should be described in this section. One example:				
Removal of sugar-sweetened beverages from all vending machines on school campuses	Study citation	Quantitative report using revenue statements for vending machines in schools	High	When replaced with water, juice, and milk, sugar-sweetened beverages can feasibly be removed from a cost perspective.

an evidence table that should be included in the evidence report for each potential intervention. The table should be divided into three parts based on the above three questions. Thus the first part of the table should summarize the evidence that answers *Why should we do something about this problem in our situation?* As described further in Chapter 5, this evidence, sometimes called needs assessment, helps the decision maker understand the scope and severity of the public health problem in the local context. The second part of the table should summarize the evidence that answers *What specifically should we do about this problem?* This evidence describes the effectiveness and impact of potential interventions and helps the decision maker select which one to implement. Finally, the third part of the table should summarize the evidence that answers *How do I implement this information for our situation?* This evidence helps the decision maker assess how a potential intervention should be implemented to achieve its desired effects in the local context. As illustrated by Table 7-2, each source of evidence and the issue it addresses should be identified, as well as the type of study design or methodology, its quality, and its major outcomes or findings. Guidance on this information can be found in Chapters 5 and 6 of this report.

Summary of Evidence

Once the potentially relevant evidence has been located and evaluated for each potential intervention as called for by the L.E.A.D. framework, it should be synthesized and summarized to help inform the decision based on the question asked. Evidence and other sources of information like theory, professional experience, local wisdom are synthesized using tools described earlier in this chapter. The following subsections address the formulation and reporting of this summary.

Why Should We Do Something About This Problem in Our Situation?

The first part of the summary should describe the local conditions to be addressed by the potential intervention. It might include, for example, the number of people affected by obesity; the rate of increase in obesity; the number of people affected by obesity-related diseases, such as type 2 diabetes, and whether those affected fall within a certain age, racial, or ethnic group more than other groups; and the associated health care costs—all specific to the particular setting.

What Specifically Should We Do About This Problem?

The second part of the summary should provide a description and rationale for the potential intervention. The synthesis of this evidence should consider the broader context for the intervention from a systems perspective, how effective the intervention might be, and its public health impact. Table 7-1, presented earlier, lists tools that can be used in assembling this evidence. Decision makers may need to balance highly effective interventions that reach only a modest number of people with less effective interventions that reach the entire population at risk. Or they may need to look at interventions for which little evidence is available on their effectiveness or impact but that enjoy community acceptance, or ones that are recommended by experts in the field based on their experience and knowledge. Guidance is offered below for summarizing the three types of evidence that can inform the selection of an intervention: (1) the broader context of the potential intervention from a systems perspective (i.e., rationale); (2) how effective the intervention is likely to be; and (3) its likely reach and impact.

The Broader Context from a Systems Perspective (i.e., rationale). The summary should clearly describe the theory used to link the outcome(s) of the intervention to obesity prevention, or the systems view taken to understand the implications of the intervention for the outcome of interest. Logic models can help predict what effect an intervention will have on behavior, whether it will lead to the desired outcome, or whether it will lead to other outcomes. Systems views highlight areas of concern and provide decision makers with options for evaluating new interventions (see Chapter 4).

Reporting Likely Effectiveness. As discussed previously, the level of certainty as to the causal relationship between an intervention and the observed outcomes and the extent to which the research results can be translated to the current situation (generalizability) both increase the likelihood that the intervention will yield favorable results. The degree of certainty needed to recommend an intervention is not absolute: interventions that are low-cost/low-risk may be recommended with less certainty than interventions that are higher-cost or higher-risk. Interventions that require substantial capital or ongoing expenditures or that have higher potential risks will need stronger evidence before being recommended (Cohen and Neumann, 2008).

In accordance with the L.E.A.D. framework, the report should present decision makers with a summary of the estimated effectiveness of the potential intervention, following the approach of the Washington State Department of Health (2005), which was subsequently cited and applied by the Texas Department of Health State Health Services (2006). In addition, to present the full spectrum of possible conclusions, the committee suggests adding a fourth category of "no likely benefit" to the Washington State and Texas approach. As a result, the committee suggests the following categories of effectiveness:

- **Effective—**One or more well-designed studies in similar contexts have shown that a potential intervention has had the desired effect on nutrition, physical activity, or obesity rate.
- **Promising—**Evidence suggests likely effectiveness in achieving intervention goals. A potential intervention is adapted from a proven intervention, with substantial alterations that (1) have a scientific basis; (2) are based on evidence from evaluation studies or reports showing meaningful, positive health or behavioral outcomes; or (3) are supported by written program evaluations that include evidence of effectiveness, formative evaluation data, and/or results consistent with theory in the case of high-reach, low-cost, replicable interventions (Brownson et al., 2009).
- **Untested—**A potential intervention has shown an effect in another context; it has a strong theoretical basis and some supporting data, but key data are missing; or the evidence shows effects on intermediate outcomes only, particularly if the link between these intermediate outcomes and obesity has not yet been firmly established.
- **No likely benefit—**Well-conducted studies fail to show a positive effect for a potential intervention, or there is evidence of harm.

An evolution has occurred in the categorical presentation of summary conclusions regarding the evidence base for public health interventions. The committee gave special weight to the classification systems used by the Washington State and Texas public health departments and noted their congruence with current academic assessment schema—particularly those of Brownson and colleagues (2009) and a paper commissioned for this report on the current review of environmental and policy interventions for childhood obesity prevention undertaken by Transtria (see Appendix B for more information on this project). Although there are differences in nuance among the various proposed schema, as noted earlier, clear communication among professionals requires the adoption of a uniform lexicon for reporting.

When possible, the summary of evidence concerning likely effectiveness should also provide an estimate of the magnitude of the effect. In the case of quantitative data, specific statistical tools will help in describing the uncertainty surrounding these estimates. In the case of qualitative data or data that are less directly applicable to the decision-making context, the range of likely results will be more difficult to describe. In this case, qualitative descriptions of prior experiences in similar settings, which are always valuable adjuncts to research results, may be the only basis for decision making.

Reporting Likely Reach and Impact. An intervention should be categorized as universal, selective, or indicated. Within the universal category, evidence for the reach and uptake of the intervention should be summarized, in quantitative terms when possible. In some cases, widespread reach may allow for significant impact from interventions that are of relatively low absolute effectiveness. An example is distributing guidance on healthful eating through the health care or education system, which may be inexpensive, involve minimal risk, and be capable of reaching a large number of community members. Presented with evidence for a sizable potential impact, decision makers may adopt interventions of this sort as part of a larger portfolio.

How Do We Implement This Information for Our Situation?

The final part of the summary should address the implementation of the potential intervention. This section, which will most commonly be in narrative form, should describe (1) the personnel and other resources deployed in the studies examined; (2) if applicable, the relationship between the size of the program staff and the population reached; (3) the institutional relationships required for implementation; and (4) any other resource requirements that would aid decision makers in understanding how best to apply the results of previous efforts to planning this intervention.

In many cases, interventions may need to be adapted to fit the local context. Community participation in this process may lead to better understanding of the local context for implementation. When sufficient data are available, realist reviews (Pawson et al., 2005) may be undertaken to guide decision makers in making adaptations to local circumstances that are likely to preserve the effectiveness of previous programs in their context.

KNOWLEDGE INTEGRATION

As noted earlier, the new knowledge obtained in following the framework needs to be incorporated into the context of the organization or system where decisions are being made, such as a public health department, a mayor's office, a health care system, or a school system. Rather than an active step in the framework, knowledge integration is the desired final outcome of the process. Knowledge integration may also generate new questions and feed back into the steps of the framework. The cyclic nature of knowledge generation, use of evidence, integration of relevant evidence into systems, and feedback is thus completed.

The concept of knowledge integration derives from the evolution of thinking about how new knowledge from research is used (Best et al., 2008b). Earlier terms such as "knowledge transfer" and "knowledge uptake" captured the need to move new knowledge generated by research into the hands of users but reflected one-way, linear thinking, with researchers simply producing new knowledge. Experience has suggested that these models do not work (Davis et al., 2003; Grimshaw et al., 2001). Other terms, such as "knowledge translation," suggest more of a relationship model incorporating the necessary interactions among individuals and the importance of social networks (Canadian Institutes of Health Research, 2008). "Knowledge integration" takes the process of translating evidence to practice a step further, indicating that relationships themselves are shaped by the organizations and systems in which they are embedded and that a system has particular dynamics, priorities, time scales, modes of communication, and expectations (Best et al., 2008a). In a similar manner, the recent emphasis of the National Institutes of Health (NIH) on dissemination and implementation research has encouraged investigators and decision makers to focus their research and programmatic questions on the integration of knowledge within systems (Kerner et al., 2005).

The concept of knowledge integration is closely linked to an appreciation for a study's generalizability, discussed throughout this report. An understanding of the context in which evidence will be used is clearly part of successful knowledge integration. The challenges of addressing a study's generalizability, such as recognizing the ecologically complex nature of communities, organizations, and health care systems as opposed to the more tightly (and artificially) controlled conditions of RCTs, have been described (Green and Glasgow, 2006).

Knowledge integration places specific emphasis on systems thinking in health sciences, which has evolved out of scientific inquiry in such fields as mathematics, biology, engineering, ecology, and management science (see Chapter 4). Examples of systems thinking come from the fields of weather forecasting, prevention of pandemics, and tobacco control (Leischow et al., 2008). In the control of obesity, multiple systems or organizations are clearly involved. The management of systems knowledge and its successful integration will therefore require an infrastructure to store, disseminate, and communicate this information to various stakeholders within a systems environment (Leischow et al., 2008) (see Chapter 10).

REFERENCES

Armstrong, R., E. Waters, L. Moore, E. Riggs, L. G. Cuervo, P. Lumbiganon, and P. Hawe. 2008. Improving the reporting of public health intervention research: Advancing TREND and CONSORT. *Journal of Public Health* 30(1):103-109.

Best, A., R. A. Hiatt, and C. D. Norman. 2008a. Knowledge integration: Conceptualizing communications in cancer control systems. *Patient Education and Counseling* 71(3):319-327.

Best, A., W. K. Trochim, J. Haggerty, G. Moor, and C. Norman. 2008b. Systems thinking for knowledge integration: New models for policy-research collaboration. In *Organizing and reorganizing: Power and change in health care organizations*, edited by E. Ferlie, P. Hyde, and L. McKee. London: Routledge.

Briss, P. A., S. Zaza, M. Pappaioanou, J. Fielding, L. Wright-De Aguero, B. I. Truman, D. P. Hopkins, P. D. Mullen, R. S. Thompson, S. H. Woolf, V. G. Carande-Kulis, L. Anderson, A. R. Hinman, D. V. McQueen, S. M. Teutsch, and J. R. Harris. 2000. Developing an evidence-based Guide to Community Preventive Services—methods. *American Journal of Preventive Medicine* 18(1, Supplement 1):35-43.

Brownson, R. C., J. E. Fielding, and C. M. Maylahn. 2009. Evidence-based public health: A fundamental concept for public health practice. *Annual Review of Public Health* 30(15):1-27.

Canadian Institutes of Health Research. 2008. *Knowledge translation—definition*. http://www.cihr-irsc.gc.ca/e/39033.html#Exchange (accessed January 4, 2010).

CDC (Centers for Disease Control and Prevention). 1996. Cigarette smoking before and after an excise tax increase and an antismoking campaign—Massachusetts, 1990-1996. *Morbidity and Mortality Weekly Report* 45(44):966-970.

Cohen, D. J., B. F. Crabtree, R. S. Etz, B. A. Balasubramanian, K. E. Donahue, L. C. Leviton, E. C. Clark, N. F. Isaacson, K. C. Stange, and L. W. Green. 2008. Fidelity versus flexibility: Translating evidence-based research into practice. *American Journal of Preventive Medicine* 35(5, Supplement):S381-S389.

Cohen, J., and P. Neumann. 2008. Using decision analysis to better evalute pediatric clinical guidelines. *Health Affairs* 27(5):146-175.

Connelly, J. B., M. J. Duaso, and G. Butler. 2007. A systematic review of controlled trials of interventions to prevent childhood obesity and overweight: A realistic synthesis of the evidence. *Public Health* 121(7):510-517.

Cullen, K. W., and K. B. Watson. 2009. The impact of the Texas public school nutrition policy on student food selection and sales in Texas. *American Journal of Public Health* 99(4):706-712.

Davis, D., M. Evans, A. Jadad, L. Perrier, D. Rath, D. Ryan, G. Sibbald, S. Straus, S. Rappolt, M. Wowk, and M. Zwarenstein. 2003. The case for knowledge translation: Shortening the journey from evidence to effect. *British Medical Journal* 327(7405):33-35.

D'Onofrio, C. N. 2001. Pooling information about prior interventions: A new program planning tool. In *Handbook of program development for health behavior research and practice*, edited by S. Sussman. Thousand Oaks, CA: Sage Publications. Pp. 158-203.

Eriksen, M. 2005. Lessons learned from public health efforts and their relevance to preventing childhood obesity. In *Preventing childhood obesity: Health in the balance*, edited by J. Koplan, C. Liverman, and V. Kraak. Washington, DC: The National Academies Press. Pp. 343-375.

Farley, T. A., A. Caffarelli, M. T. Bassett, L. Silver, and T. R. Frieden. 2009. New York City's fight over calorie labeling. *Health Affairs* 28(6):w1098-w1109.

Glasgow, R. E., L. M. Klesges, D. A. Dzewaltowski, P. A. Estabrooks, and T. M. Vogt. 2006. Evaluating the impact of health promotion programs: Using the RE-AIM framework to form summary measures for decision making involving complex issues. *Health Education Research* 21(5):688-694.

Glasziou, P., and H. Longbottom. 1999. Evidence-based public health practice. *Australian and New Zealand Journal of Public Health* 23(4):436-440.

Green, L. W., and R. E. Glasgow. 2006. Evaluating the relevance, generalization, and applicability of research: Issues in external validation and translation methodology. *Evaluation & the Health Professions* 29(1):126-153.

Green, L. W., and M. W. Kreuter. 2005. *Health program planning: An educational and ecological approach*. 4th ed. New York: McGraw-Hill.

Greenhalgh, T., E. Kristjansson, and V. Robinson. 2007. Realist review to understand the efficacy of school feeding programmes. *British Medical Journal* 335(7626):858-861.

Grimshaw, J. M., L. Shirran, R. Thomas, G. Mowatt, C. Fraser, L. Bero, R. Grilli, E. Harvey, A. Oxman, and M. A. O'Brien. 2001. Changing provider behavior: An overview of systematic reviews of interventions. *Medical Care* 39(8, Supplement 2):II-2-II-45.

Guyatt, G. H., A. D. Oxman, R. Kunz, R. Jaeschke, M. Helfand, A. Liberati, G. E. Vist, and H. J. Schunemann. 2008a. Incorporating considerations of resources use into grading recommendations. *British Medical Journal* 336(7654):1170-1173.

Guyatt, G. H., A. D. Oxman, G. E. Vist, R. Kunz, Y. Falck-Ytter, P. Alonso-Coello, and H. J. Schunemann. 2008b. GRADE: An emerging consensus on rating quality of evidence and strength of recommendations. *British Medical Journal* 336(7650):924-926.

Hawe, P., A. Shiell, and T. Riley. 2004. Complex interventions: How "out of control" can a randomised controlled trial be? *British Medical Journal* 328(7455):1561-1563.

Health Canada. 2000. *Health Canada decision-making frameowrk for identifying, assessing, and managing health risks*. Ontario: Health Canada.

IOM (Institute of Medicine). 2007. *Nutrition standards for schools: Leading the way toward healthier youth*. Edited by V. A. Stallings and A. L. Yaktine. Washington, DC: The National Academies Press.

IOM. 2009. *Local government actions to prevent childhood obesity.* Edited by L. Parker, A. C. Burns, and E. Sanchez. Washington, DC: The National Academies Press.

Jacobs, F. 2003. Child and family program evaluation: Learning to enjoy complexity. *Applied Developmental Science* 7(2):62-75.

Jilcott, S., A. Ammerman, J. Sommers, and R. E. Glasgow. 2007. Applying the RE-AIM framework to assess the public health impact of policy change. *Annals of Behavioral Medicine* 34(2):105-114.

Kaiser Permanente Institute for Health Research. 2010. *What is RE-AIM?* http://www.re-aim.org/whatisREAIM/index.html (accessed March 9, 2010).

Kerner, J., B. Rimer, and K. Emmons. 2005. Introduction to the special section on dissemination—dissemination research and research dissemination: How can we close the gap? *Health Psychology* 24(5):443-446.

Kreuter, M. W., C. De Rosa, E. H. Howze, and G. T. Baldwin. 2004. Understanding wicked problems: A key to advancing environmental health promotion. *Health Education and Behavior* 31(4):441-454.

Leischow, S. J., A. Best, W. M. Trochim, P. I. Clark, R. S. Gallagher, S. E. Marcus, and E. Matthews. 2008. Systems thinking to improve the public's health. *American Journal of Preventive Medicine* 35(2, Supplement 1):S196-S203.

Martin, J. B., L. W. Green, and A. C. Gielen. 2007. Potential lessons from public health and health promotion for the prevention of child abuse. *Journal of Prevention and Intervention in the Community* 34(1-2):205-222.

Mercer, S. L., L. K. Kanh, L. W. Green, A. C. Rosenthal, R. Nathan, C. G. Husten, and W. H. Dietz. 2005. Drawing possible lessons for obesity prevention and control from the tobacco control experience. In *Obesity prevention and public health*, edited by D. Crawford and R. W. Jeffery. London: Oxford University Press, Inc. Pp. 231-251.

Montori, V. M., and G. H. Guyatt. 2001. Intention-to-treat principle. *Canadian Medical Association Journal* 165(10):1339-1341.

Pawson, R., T. Greenhalgh, G. Harvey, and K. Walshe. 2005. Realist review—a new method of systematic review designed for complex policy interventions. *Journal of Health Services Research and Policy* 10(Supplement 1):21-34.

Raczynski, J. M., J. W. Thompson, M. M. Phillips, K. W. Ryan, and H. W. Cleveland. 2009. Arkansas Act 1220 of 2003 to reduce childhood obesity: Its implementation and impact on child and adolescent body mass index. *Journal of Public Health Policy* 30(Supplement 1): S124-S140.

Rychetnik, L., P. Hawe, E. Waters, A. Barratt, and M. Frommer. 2004. A glossary for evidence based public health. *Journal of Epidemiology and Community Health* 58(7):538-545.

Sege, R. D., and E. De Vos. 2008. Care for children and evidence-based medicine. *Pediatric Annals* 37(3):168-172.

Shrier, I., J. F. Boivin, R. J. Steele, R. W. Platt, A. Furlan, R. Kakuma, J. Brophy, and M. Rossignol. 2007. Should meta-analyses of interventions include observational studies in addition to randomized controlled trials? A critical examination of underlying principles. *American Journal of Epidemiology* 166(10):1203-1209.

Siegel, M., P. D. Mowery, T. P. Pechacek, W. J. Strauss, M. W. Schooley, R. K. Merritt, T. E. Novotny, G. A. Giovino, and M. P. Eriksen. 2000. Trends in adult cigarette smoking in California compared with the rest of the United States, 1978-1994. *American Journal of Public Health* 90(3):372-379.

Swinburn, B., T. Gill, and S. Kumanyika. 2005. Obesity prevention: A proposed framework for translating evidence into action. *Obesity Reviews* 6(1):23-33.

Texas Department of State Health Services. 2006. *Texas obesity policy portfolio*. Austin, TX: Center for Policy and Innovation.

UK Government Foresight Programme. 2007. *Tackling obesities: Future choices—qualitative modelling of policy options*. London, UK: Crown.

Washington State Department of Health. 2005. *Nutrition and physical activity: A policy resource guide*. http://www.dshs.state.tx.us/obesity/pdf/WAnutritionphysicalactivityguide.pdf (accessed December 1, 2009).

8

Opportunities to Generate Evidence

KEY MESSAGES

- The evidence base to inform decisions about obesity prevention is extremely limited, even taking into account the expanded view of evidence proposed in this report (Chapter 5). Each step in the L.E.A.D. framework can provide important insights about the types of evidence that are needed.
- In obesity prevention–related research, the evaluation of ongoing and emerging initiatives is of particular priority.
- Future research on obesity prevention, and in public health more generally, can employ a broad array of study designs that support valid inferences about the effects of policies and programs without experimentation, promoting transdisciplinary exchange.
- Published peer-reviewed reports on the results of obesity prevention efforts often lack useful information related to generalizability to other individuals, settings, contexts, and time frames, adding to the problem of incomplete evidence for decision making.
- If obesity prevention actions must be taken when evidence is limited, the L.E.A.D. framework calls for developing credible evidence about those actions for use in decision making about future efforts, including the use of natural experiments, and emerging and ongoing programs through "evaluability assessments" and continuous quality assessments.

The preceding chapter outlines a standard for assembling evidence that requires access to potentially useful sources of information; the blending of theory, expert opinion, experience, and local wisdom; the availability of scientifically trained staff or colleagues skilled in using these resources; a time window that allows for the process of locating, evaluating, and assembling evidence; and, most important, the existence of relevant evidence. Use of the L.E.A.D. framework to broaden what is considered useful, high-quality evidence and to gradually increase the availability of such evidence will help in attaining this standard. Together with the importance of taking a systems approach and making the best possible use of diverse types of evidence that are relevant to the user's perspective and the question being asked, a major emphasis of this report is the urgent need to generate more evidence to inform efforts to address obesity prevention and other complex public health problems. Much of what is called for in the framework will meet the expectations and needs of decision makers when sufficient evidence exists, although clearly such is not yet the case (see Chapter 3). On the other hand, failure to find relevant evidence may either (erroneously) reinforce the perception that effective interventions cannot be identified, or increase skepticism or resistance on the part of decision makers—who must proceed in the interim—with respect to the utility of incorporating evidence into their decision-making process. As described in Chapter 7, when decision makers face choices that must be made and actions that must be taken in the relative absence of evidence, or at least on the basis of inconclusive, inconsistent, or incomplete evidence, the L.E.A.D. framework calls for critical evaluation and systematic building on experience in a continuous translation, action, monitoring, surveillance, and feedback loop (i.e., matching, mapping, pooling, and patching; see Appendix E for more detail).

The purpose of this chapter is to motivate researchers and others whose primary role is to generate (i.e., support, fund, publish) evidence to adopt the L.E.A.D. framework as a guide to identifying and generating the types of evidence that are needed to support decision making on obesity prevention and other complex public health problems (see Figure 8-1). The focus is on evidence related to "What" and "How" questions, that is, evidence demonstrating the types of interventions that are effective in various contexts; what their impacts are; and which are relatively more or less effective, whether they are associated with unexpected benefits or problems, and what practical issues are involved in their implementation. As noted in Chapter 3 and defined in Chapter 5, these concerns point to areas in which a lack of evidence is most likely to be problematic. This chapter also anticipates a cycle that begins with planning from incomplete evidence, blended with theory, expert opinion, experience, and local wisdom (see Chapter 7), and ends with evaluating the consequences of interventions. Such a cycle, in turn, may produce the most credible evidence for other jurisdictions seeking practical models to emulate.

The chapter begins by briefly reviewing existing evidence needs and outlining the need for new directions in evidence generation and transdisciplinary exchange. It then addresses the limitations in the way evidence is reported in scientific journals and the need to take advantage of natural experiments and emerging and ongoing interventions as sources of practice-based evidence to fill the gaps in the best available evidence. The chapter concludes with a discussion of alternatives to randomized

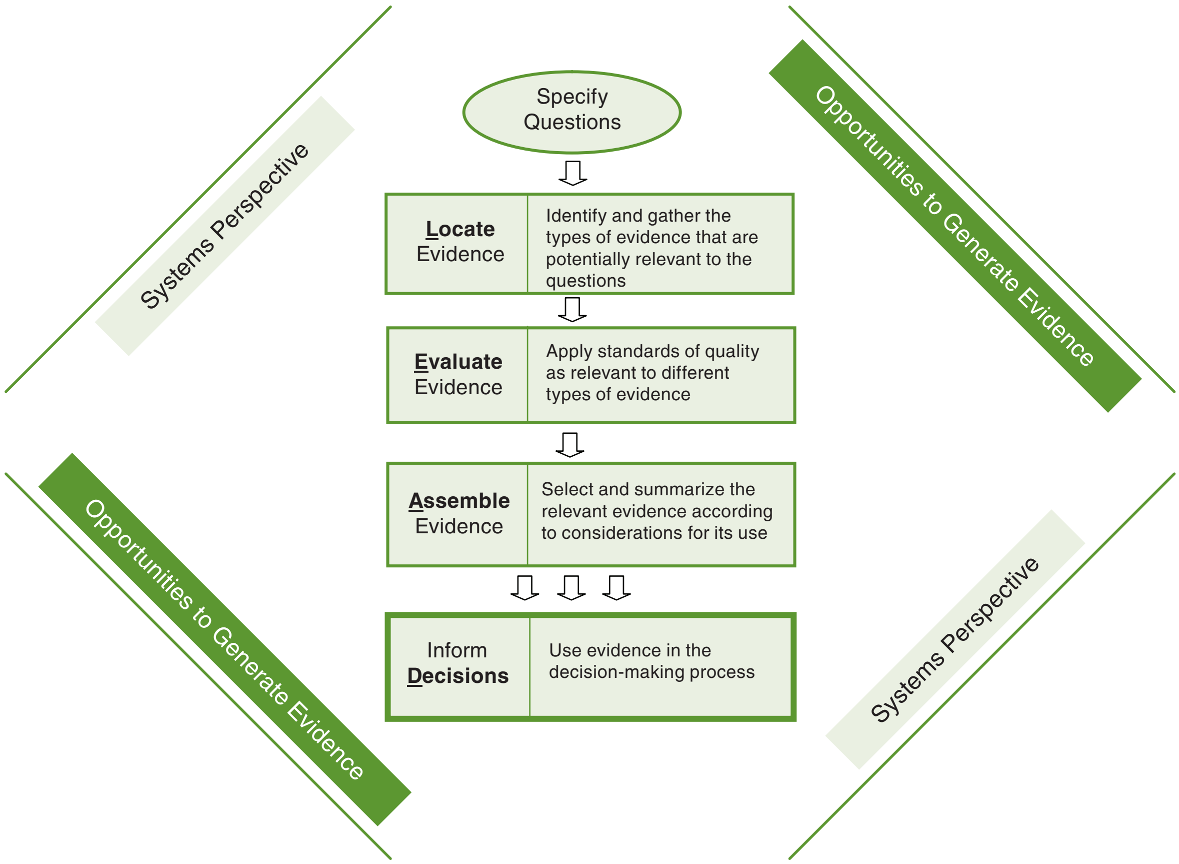


FIGURE 8-1 The **L**ocate Evidence, **E**valuate Evidence, **A**ssemble Evidence, Inform **D**ecisions (L.E.A.D.) framework for obesity prevention decision making.
NOTE: The element of the framework addressed in this chapter is highlighted.

experiments, with a focus on the level of certainty of the causal relationship between an intervention and the observed outcomes.

EXISTING EVIDENCE NEEDS

The generation of evidence related to interventions and their effectiveness can be approached through evaluation research, where evaluation is defined as "a systematic assessment of the quality and effectiveness of an initiative, program, or policy . . ." (IOM, 2007, p. 26). As explained in Chapter 1, the 2007 Institute of Medicine (IOM) report *Progress in Preventing Childhood Obesity*, which is current through 2006, strongly emphasizes the need for evaluation of ongoing and emerging initiatives. The committee that produced the report found that, in response to the urgency of the problem of childhood obesity with respect to prevalence and economic costs, numerous efforts were being undertaken on the basis of what was already known from theory or practice. Given the inherent relevance of implemented programs to natural settings, such spontaneous or endogenous interventions may be of most interest to decision makers. However, many of these programs are not evaluated. Issues for consideration in increasing the use and appropriateness of such evaluation, taken from the 2007 IOM committee's assessment, are shown in Box 8-1.

A working group convened by the National Heart, Lung, and Blood Institute of the National Institutes of Health (NIH) also has made several recommendations for research needs related to the prevention of child obesity, many of which are directly relevant to the types of population-based approaches addressed in this report (Pratt et al., 2008). Several of these recommendations are listed in Box 8-2. Consistent with the issues highlighted by the 2007 IOM committee, the NIH recommendations include the evaluation of existing promising programs, as well as the conduct of studies of multilevel and multicomponent interventions.

The examples in Box 8-2 relate specifically to child obesity. However, the need for more evidence relevant to decision making is also recognized with respect to obesity in adults and other complex public health problems. The Congressional Budget Office (CBO), for example, has identified several areas in which having more evidence would be helpful. CBO found some evidence on net cost reductions for certain disease management programs, but it was unclear whether these strategies could be replicated or applied in a broader population. In general, CBO found the availability of clinical and economic research assessing and comparing treatments for preventive services to be limited (CBO, 2008).

Box 8-1
Considerations for Increasing Evaluation of Obesity Prevention Initiatives

- Evaluation is often not a priority for individuals and organizations that are developing a new policy, program, or intervention.
- The time and resources needed to ensure appropriate baseline and outcome measures may not be available or allocated.
- Evaluation may seem too technically complex.
- Evaluation may not be identified as a responsibility of those undertaking the initiative.
- Evaluation outcomes are not always matched to the nature and stage of the program; e.g., intermediate outcomes such as change in food intake, physical activity, or TV viewing may be more appropriate than BMI (body mass index) for short-term programs.
- Evaluations of multiple, linked programs require collaborative efforts and systems approaches.

SOURCE: IOM, 2007, pp. 27-28.

Box 8-2
Selected Recommendations for Research in Childhood Obesity Prevention

Settings

- Test interventions with physicians and other health care providers, combined with community involvement (e.g., train physicians to screen, nurses to be coaches, and health care settings to refer to community resources) (for young children).
- Conduct interventions in a variety of settings (e.g., home; child care; the U.S. Department of Agriculture's Special Supplemental Nutrition Program for Women, Infants, and Children [WIC]; and health care settings) (for young children).
- Examine multilevel and multicomponent, community-based interventions in multiple settings (e.g., schools, health care, home, community, built environment, public policy, social marketing, diet, physical activity behaviors).
- Test a multilevel, comprehensive intervention that targets minority and low-income populations (e.g., culturally appropriate ways to reach Latino, African-American, Native American, and Asian/Pacific Islander children).
- Develop and test interventions that can be incorporated effectively into existing school and community infrastructures (e.g., curriculum, physical activity, school lunch programs) to maximize effectiveness and minimize cost.
- Conduct intervention studies that address issues related to the interface between individual behaviors and the environment.

Implementation, Dissemination, Translation, Evaluation

- Identify and test approaches for community partnerships in the dissemination and implementation of evidence-based obesity-prevention programs.
- Evaluate the effectiveness of existing promising programs.

Methodology

- Support methodological research on study designs and analytic approaches (identify optimal study designs and analytic approaches for various types of research questions).
- Use appropriate study designs and methods, including natural experiments, quasi-experimental designs, and randomized designs; develop time-sensitive funding mechanisms for natural experiments.

High-Risk Populations

- Study a diversity of high-risk and understudied subgroups, including low-income families, ethnically and socioeconomically diverse populations, boys, and children in rural communities, as well as immigrants.
- Conduct environmental and policy intervention research to improve access to healthy foods and the opportunity for physical activity in low-income communities.

Other Recommendations

- Analyze effectiveness–intervention studies for their cost-effectiveness.

SOURCE: Adapted and reprinted from Pratt et al., Copyright © 2008, with permission from Elsevier.

THE NEED FOR NEW DIRECTIONS AND TRANSDISCIPLINARY EXCHANGE

The two methodological recommendations in Box 8-2 highlight one of the major problems encountered by those who wish to generate policy- or program-relevant evidence regarding population health problems such as obesity. These recommendations refer to the need to identify optimal and appropriate study designs, including designs that will allow for timely assessment of initiatives that constitute "natural experiments" (detailed later in this chapter); quasi-experimental designs are mentioned in addition to randomized designs. One major question raised by these recommendations relates to the prevailing preference for randomized controlled trials (RCTs) in biomedical science because of their advantages for drawing causal inferences: Are there good alternatives to randomized designs that will accomplish the same thing but can be implemented more flexibly in natural or field settings? (West et al., 2008). A related question, also reflected in the considerations from the 2007 IOM report and the NIH recommendations, is the complexity of many interventions conducted in the community: To what degree do the multiple components of interventions need to be separated to isolate their causal effects? In some, perhaps many, cases, the separation may be artificial—the effectiveness of each component may be dependent on the others, and the full treatment package may be of primary interest (West and Aiken, 1997). In other cases, components may need to be deleted because of ineffectiveness or iatrogenic effects or for cost-effectiveness reasons. Chapter 4 provides a fuller discussion of these conceptual issues.

A more general issue is how experts who are trained in a particular research approach can expand their capabilities for addressing issues that can be well studied only with different, less familiar approaches. For example, how do biomedical researchers, who have traditionally conducted the research on obesity, become conversant with methods for evaluating evidence that may be available from other fields and with the experts, or expertise, in implementing those methods? How do researchers in fields other than biomedicine—for example, education, community design, or economics—become involved and expert in studying problems, such as obesity, some aspects of which do not lend themselves to their typical methods? And how can obesity researchers benefit from the scholarship of colleagues who have focused on different, similarly complex public health problems, such as tobacco control?

Transdisciplinary exchange refers to researchers' use of a "shared conceptual framework drawing together disciplinary-specific theories, concepts, and approaches to address a common problem" (Rosenfield, 1992, p. 1351). Both the medical and social sciences are challenged to expand their notion of methods and study designs that can inform the study of obesity prevention and other population health problems by considering contextual concerns (social, economic, environmental, institutional, and political factors) that influence health outcomes. For example, NIH's National Institute on Aging has supported a network or team approach to studying the "biological, psychological and social pathways to positive and not-so-positive health" (Kessel and Rosenfield, 2008, p. S229); the result has been enhanced quality of published research that crosses disciplinary lines. Funding agencies such as NIH have therefore been encouraged to continue to support this type of research and urge researchers to value collaboration and partnerships in a variety of fields.

Stimulated by a 2006 National Cancer Institute conference that highlighted the need for transdisciplinary research in health, Kessel and Rosenfield (2008) describe a number of factors that have both facilitated and constrained transdisciplinary science (Table 8-1). The authors identified these factors from research programs that have successfully crossed disciplinary boundaries. Reflecting the type of transdisciplinary

TABLE 8-1 Factors Facilitating and Constraining Transdisciplinary Team Science

Factor	Facilitating	Constraining
Focus on major problems	PIs able to bring researchers together across disciplines and program-unifying themes	Some areas seen as unrealistic Lack of integrative research framework Few "how-to" models
Team members	Possess complementary and intersecting skills Able to develop common language Positive open attitude Appreciative of others' knowledge Shared understanding of scientific problem Mutual trust and respect Open to mentoring	See skills as competitive Tension between solo and collaborative work Power–prestige differences social and medical sciences Worry about diffusion of focus and loss of identity Research seen as time-consuming/multiple projects Disincentive for practitioners Sharing credit affects promotion, tenure, publications, funding
Training	Complementary training Mentored graduate students to participate in transdisciplinary research team SERCA grants for training in new field	Historical barriers across fields Location of departments Funding limited
Institutions	Support, promote, and fund centers, networks, and teams across disciplines, departments, and medical and social science facilities on same campus	Rigid university policies Centers lack funds
Technology	Facilitate communication even when teams and researchers physically dispersed	
Funding	Foundations and government support network/team approach (e.g., MacArthur, NIH)	Grant applications more challenging, time-consuming
Publication		Journals discourage multiple authors Peer review hard to judge Need to frame more narrowly

NOTE: NIH = National Institutes of Health; PI = principal investigator; SERCA = Special Emphasis Research Career Award.
SOURCE: Reprinted from Kessel and Rosenfield, Copyright © 2008, with permission of Elsevier; and Kessel and Rosenfield, page 401—Table 19.3, "Fostering interdisciplinary innovation: The way forward" from "Expanding the boundaries of health and social science: Case studies of interdisciplinary innovation" (2003), by permission of Oxford University Press, Inc.

exchange that is needed among researchers, the final section of this chapter examines alternatives to randomized experiments from the perspective of evaluation experts in the behavioral sciences, economometrics, education, public health, and statistics and explains why these alternatives are particularly important for obesity prevention (a more detailed discussion of selected study designs is offered in Appendix E).

LIMITATIONS IN THE WAY EVIDENCE IS REPORTED IN SCIENTIFIC JOURNALS

Decision makers (e.g., policy makers, professional caregivers, public health officials, and advocates) have concerns beyond scientists' certainty of causal relationships in judging the utility and persuasiveness of evidence (see Chapter 7). All of these concerns fall to some degree under the broad rubric of generalizability (Garfield et al., 2003; Glasgow et al., 2003, 2006; Green, 2001; Green and Glasgow, 2006; Green and Lewis, 1981; Green et al., 1980; Shadish et al., 2002, Chapters 11-13). As described in Chapter 5, the generalizability of a study refers to the degree to which its results can be expected to apply equally to other individuals, settings, contexts, and time frames. The generalizability of evidence-based practices is seldom considered in individual studies or systematic reviews of the evidence, but is often decisive in decision makers' adoption of research evidence for practical purposes such as obesity prevention. Some authors have begun to make a regular practice of noting the "applicability" of findings, with caveats when the range of populations or settings in which the evidence from trials was derived is notably limited (Shepherd and Moore, 2008). Others have proposed combining reviews of clinical and community evidence in a more "ecologically comprehensive" (multimethod, multilevel) approach to the use of evidence in such areas as tobacco and obesity control (Ockene et al., 2007). And some journals have begun to make the generalizability of a study more of an issue in considering manuscripts (e.g., Eriksen, 2006; Steckler and McLeroy, 2008), on occasion with qualifications and noted constraints (Patrick et al., 2008).

The usual manner of reporting results of obesity prevention efforts in journals often adds to the problem of incomplete evidence because useful aspects of an intervention and research related to its generalizability are not discussed. At a meeting of the editors of 13 medical and health journals, several agreed to make individual and joint efforts to devote more attention to issues of a study's generalizability and to press for more reporting of (1) recruitment and selection procedures, participation rates, and representativeness at the levels of individuals, intervention staff, and delivery settings; (2) the level of consistency with which the intervention being tested was implemented in the study; (3) the impact on a variety of outcomes, especially those important to patients, clinicians, program cost, and adverse consequences; and (4) in follow-up reports, attrition of subjects in the study at all levels, long-term effects on outcomes, and program institutionalization, modification, or discontinuation (Green et al., 2009; supported by Cook and Campbell, 1979, pp. 74-80; Glasgow et al., 2007; Shadish et al., 2002). These moves by journal editors relevant to obesity control hold promise for the greater relevance and usefulness of the published scientific literature for other decision makers, responding to some of their major concerns.

Specifically in childhood obesity prevention, Klesges and colleagues (2008) examined studies published between 1980 and 2004 that were controlled, long-term research trials with a behavioral target of either physical activity or healthful eating or both, together with at least one anthropometric outcome. Using review criteria for a study's generalizability to other individuals, settings, contexts, and time frames (external validity) developed by Green and Glasgow (2006), they found that all of the 19 publications that met their selection criteria lacked full reporting on the 24 dimensions of external validity expected in an optimal paper to enable users to judge potential generalizability (see Box 8-3). Median reporting over all elements was 34.5 percent; the mode was 0 percent with a range of 0 percent to 100 percent. Only four dimensions (descriptions of the target audience and target setting, inclusion–exclusion criteria, and attrition rate) were reported in at least 90 percent of the studies. Most infrequent were reports of setting-level selection criteria and representativeness, characteristics of intervention staff, implementation of intervention content, costs, and program sustainability. These limitations of individual studies are also seen and sometimes multiplied in systematic reviews, such as meta-analyses, of whole bodies of literature. The cumulative problems of inadequate reporting of sampling, settings, and interventions have been noted, for example, in meta-analyses of the patient education literature on preventive health interventions in clinical settings (Simons-Morton et al., 1992; Tabak et al., 1991).

These findings provide strong support for the conclusion of Klesges and colleagues (2008) that the aspects of generalizability that potential users need most to see reported more thoroughly in the published evidence are the "3 Rs": the *representativeness* of participants, settings, and intervention staff; the *robustness* of the intervention across varied populations and staffing or delivery approaches; and the *replicability* of study results in other places. The specific questions most decision makers will have within these broad categories relate to cost (affordability); scalability; and acceptability in particular populations, times, and settings.

Even with more complete reporting on these issues of a study's generalizability to other populations, gaps will inevitably remain. RCTs can never fill all of the cells in a matrix of potentially relevant evidence representing all combinations of a study's dimensions of generalizability: population × setting × intervention × time × staffing × other resources. The empty cells in such a matrix require potential users of evidence to make inferential leaps or more studied extrapolations from the existing coverage of the evidence to their own population, setting, intervention, time, staffing, and other resources. In short, users should bring to bear on their decisions their own theories or assumptions about the fit of the evidence to their situation, which will vary along each of the above dimensions.

A particular challenge for obesity prevention, as in some other areas of chronic disease control, is the multiplicity and complexity of these dimensions. For each

Box 8-3
Percentage of 19 Studies Reporting External Validity[a] Dimensions, 1980-2004

External Validity Dimension[b]	Percent Reporting
Reach and Representativeness	
Individual Participants	
Description of targets	100
Individual inclusion/exclusion criteria	90
Participation rate	63
Representativeness	10
Setting	
Description of targets	95
Setting inclusion/exclusion criteria	11
Participation rate	22
Representativeness	0
Delivery Staff	
Participation rate	5
Implementation and Adaptation	
Consistency of implementation	26
Staff expertise or training	89
Variations in implementation by staff	5
Program adaptation	42
Outcomes of Decision Making	
Outcomes compared with goal	37
Adverse consequences	32
Moderation of effect by participant characteristic(s)	53
Moderation of effect by staff/setting	10
Number of sessions or time needed to deliver intervention	68
Costs	0
Maintenance and Institutionalization	
Long-term effects (at least 12 months)	74
Program sustainability	0
Attrition rate	100
Differential attrition by condition tested	21
Drop-out representativeness	42

[a] External validity is defined according to Leviton (2001).

[b] See Green and Glasgow (2006) for a detailed description of coding dimensions.

SOURCE: Reprinted from Klesges et al., Copyright © 2008, with permission from Elsevier.

intervention, for example, there will be a multiplicity of points of intervention, from individuals to families, social groups, organizations, whole communities, regions, or states, each calling for different sources or types of evidence. Complexity will further confound the translation of evidence with respect to the projected time frame, even across the lifespan, to account for variations in interventions that make them age-appropriate and account for such relevant background variables as media coverage of the issue, social discourse, and changing social norms.

WAYS TO FILL THE GAPS IN THE BEST AVAILABLE EVIDENCE

This section examines the need to take advantage of natural experiments and emerging and ongoing interventions as sources of practice-based evidence to compensate for the fundamental limitations of even the best available evidence.

Natural Experiments as Sources of Practice-Based Evidence

Rather than waiting for the funding, vetting, implementation, and publication of RCTs (and other forms of research) to answer practical and locally specific questions, one alternative is to treat the myriad programs and policies being implemented across the country as natural experiments (Ramanathan et al., 2008). Applying more systematic evaluation to interventions as they emerge, even with limited experimental control over their implementation, will yield more immediate and practice-based evidence for what is possible, acceptable, and effective in real-world settings and populations. What makes such evaluation of these natural experiments even more valuable is that it produces data from settings that decision makers in other jurisdictions can view as more like their own than the settings of the typical published trials. Seeing how other state and local jurisdictions are performing in a given sphere of public concern also may activate competitive instincts that spur state and local decision makers to take action.

The Centers for Disease Control and Prevention's (CDC's) Office on Smoking and Health used this strategy of making comparisons across jurisdictions when California and Massachusetts raised cigarette taxes and launched tax-based, comprehensive statewide tobacco control programs (see Chapter 4). Noting the accelerated rates of reduced tobacco consumption in those states compared with the other 48 states, CDC collaborated with the two states to evaluate these natural experiments and analyze their components (e.g., mass media, school programs) and the associated costs. CDC then offered the per capita costs of each component as a recommended basis for budgeting for other states wishing to emulate these successful programs (CDC, 1999). Many of the states used these budgetary allocations, at least temporarily before the tobacco settlement funds from the tobacco industry were diverted to general revenue or other purposes.

"Evaluability Assessment" of Emerging Innovations

As experience accumulates with local initiatives and innovative practices that have not yet been subjected to systematic evaluation, that experience can be disseminated to the larger community of public health officials. Funding should be directed at institutionalizing methods and resources that support a culture of local experimentation, evaluation, and continuous quality assessment and sharing. One such method that serves the broader community of decision makers is called "evaluability assessment" (Leviton et al., 2010a) or the systematic screening and assessment (Leviton et al., 2010b). This method was originally developed to assess whether a program had been sufficiently implemented before an outcome evaluation was undertaken (Wholey, 1977). It has been defined more recently as assessing the rationale, design, implementation, and characteristics of a program before a full-scale evaluation is undertaken to be sure the evaluation will be feasible and useful (Leviton et al., 1998; Wholey, 1994).

Recently, a Robert Wood Johnson Foundation/CDC partnership conducted a major evaluability assessment that involved taking an inventory of purported innovative programs and policies in childhood obesity control, assessing their potential for evaluation, and providing resources to assist in their evaluation (RWJF, 2006, 2009). The partners believed an evaluability assessment approach to the evidence in childhood obesity control was necessary because so many formally untested ideas, policies, and strategies were emerging across the country. Decision makers had taken seriously the public health warning signals of the obesity epidemic, but found little in the scientific literature on tested policies and methods for addressing the problem in their communities. They had no choice but to innovate, and the Foundation and CDC, also facing the paucity of evidence, had no choice but to seize upon the strands of practice-based evidence these innovations might offer. They conducted a national survey of experts and observers in the field, inviting nominations of innovative policies and programs. From the pool of hundreds, with the aid of a national panel of experts, they systematically assessed the characteristics of each that would qualify it as promising in its reach, effectiveness, adoption potential, and sustainability. They then conducted site visits to those deemed promising and from these site visits produced logic models of the mediating variables addressed by each and the evaluation capacity of the staff. From these assessments, they selected the interventions that were most worthy of formal evaluation.

This process of a centralized inventory of and evaluation assistance to the most promising interventions warrants particular attention when the urgency of action is great and the evidence for specific actions is thin. Local program staff will often need assistance with design, measurement, conduct, and analysis issues when carrying out an evaluation (Green, 2007). Without such assistance, many evaluations might respond to some local needs but would not provide the practice-based evidence of greatest use to others. The main lesson from experience with evaluability assessment is the advantage of the bottom-up, practice-driven approach over the top-down, science-driven process for generating evidence for problem solving. Taking full advantage of this lesson will require building an infrastructure of surveillance capacity and tools for self-monitoring of the implementation and effects of interventions.

Continuous Quality Assessment of Ongoing Programs

Once interventions have been established as evidence-based (in the larger sense suggested in this chapter of a combination of evidence, theory, professional experience, and local participation), continuous assessment of the quality of implementation is necessary to ensure that the interventions do not slip into a pattern of compromised effort and attenuated resources and that they build on experience with changing circumstances, clientele, and personnel. Methods for continuous quality assessment have been borrowed from industry (e.g., Edwards and Schwarzenberg, 2009) and adapted to public health program assessment to support the development of practice-based evidence in real time with real programs (Green, 2007; Katz, 2009; Kottke et al., 2008).

Continuous quality assessments of individual interventions can be pooled and analyzed for their implications for the adjustment of programs and policies to changing circumstances. These findings, in turn, can be pooled and analyzed at the state and national levels to derive guidelines, manuals, and interactive online guides for a combination of best practices and best processes. What meta-analyses and other systematic reviews may be telling us, with their inconsistent results or variability of findings over time, is that there is nothing inherently superior about most intervention practices. Rather, social, epidemiological, behavioral, educational, environmental, and administrative diagnoses lead to the appropriate application of an intervention to suit a particular purpose, population, time, setting, and other contextual variables. Such was the conclusion of a comparative meta-analysis of patient education and counseling interventions (Mullen et al., 1985).

Similarly, in clinical practice, where the application of evidence-based medicine (EBM) is generally expected, "Physicians reported that when making clinical decisions, they more often rely on clinical experience, the opinions of colleagues, and EBM summarizing electronic clinical resources rather than refer directly to EBM literature" (Hay et al., 2008, p. 707). These alternatives to the direct and simple translation of EBM guidelines, as in other fields of practice, should not be surprising in the face of the limited representation of types of patients or populations and circumstances in EBM studies. The challenge, then, is how to use the experience and results from these combinations of explicit scientific evidence and tacit experiential evidence to enrich the evidence pool. Hay and colleagues (2008) suggest methods of "evidence farming" to systematize the collection of data from clinical experience to feed back into the evidence–practice–evidence cycle.

ALTERNATIVES TO RANDOMIZED EXPERIMENTS

As discussed in previous chapters, randomized experiments are generally viewed as the gold standard for research designs. A prototype is a RCT of a drug treatment compared with a placebo or other drug. In a prevention context, an RCT would be a study in a community or broader setting in which individuals or groups would be assigned by the researchers to experience or be exposed to different interventions. When a randomized experiment is properly implemented and its assumptions can be met, it produces results that are unrivaled in the certainty with which they support causal inferences in the specific research context in which the trial was conducted. According to Shadish and colleagues (2002), the advantages of RCTs are that they protect against threats to the level of certainty of the relationship between an intervention and the observed outcomes due to history, maturation, selection, testing and instrumentation biases, ambiguous temporal precedence, and the tendency of measurement to regress to the mean. As discussed here, however, the RCT has several limitations, some inherent and some associated with typical implementations, that can decrease its value as a tool for informing decision makers addressing complex public health problems. (See also Mercer and colleagues [2007], who summarize discussions at a symposium of experts that weighed the strengths, limitations, and trade-offs of alternative designs for studying the effectiveness and translation of complex, multilevel health interventions.)

RCTs require that a pool of subjects be identified who are willing to be randomized and that providers be available who are willing to deliver the identical treatment or control intervention to each subject according to the randomization protocol. Randomization has typically been conducted at the level of individual participants. It may also be conducted with interventions being delivered to groups or communities (termed "cluster randomization"; see Donner and Klar, 2000; Kriemler et al., 2010; Murray, 1998). Or it may be conducted at multiple sites, often in different geographic regions, using parallel RCTs in which individual participants or groups are assigned to the treatment or control intervention at each site, with the results being combined (see Raudenbush and Liu, 2000). Cluster randomization designs offer new possibilities to study new types of interventions, potentially with different interventions occurring at multiple levels (e.g., individual, group, small community). Multisite designs offer possibilities to study the effects of interventions with more diverse populations, settings, and even sets of treatment providers. At the same time, cluster and multisite randomized trials raise new challenges for design and implementation, as well as statistical analysis and interpretation (e.g., Varnell et al., 2004).

With all of these types of randomized trials, moreover, practical problems commonly arise when structural or policy interventions are being studied (Bonell et al., 2006). Subjects may not want to be randomized; randomization may not be accepted, or may not be ethical or practical in the research context; or only atypical participants may be willing to be randomized.

Consider, for example, the difficulty of using an RCT design to study the effects on health of secondhand tobacco smoke or the removal of physical education classes from schools. Both of these exposures are unlikely to be within the control of a researcher. In other cases, only atypical participants would be willing to participate. An example is an RCT of a faith-based or spiritually oriented wellness group compared with a non-faith-based group because people cannot be assigned to be or not be spiritually oriented. Highly religious participants might refuse to be assigned to a non-faith-based program, while nonreligious participants might be unable to participate sincerely in a faith-based program. Or the faith-based treatment providers might believe that all people desiring the intervention should receive it. A cardiovascular disease prevention study among African American faith-based organizations in Baltimore, Maryland, encountered both of these challenges in a randomized trial (Yanek et al., 2001). Many pastors of potentially participating churches were uncomfortable with randomization, wanting to know their intervention assignment before agreeing to participate. In addition, an intended comparison between a spiritually oriented and standard version of the program could not be implemented because those in the standard condition spontaneously incorporated spirituality. These issues decrease the likelihood that an RCT can be implemented successfully and potentially decrease the likelihood that its results can ultimately be generalized to important policy contexts (Green and Glasgow, 2006; Shadish and Cook, 2009; Weisz et al., 1992).

In addition, the need to identify a pool of subjects willing to be randomized subtley privileges certain types of interventions. RCTs involving interventions designed to influence individuals or small groups (cluster randomization) are far easier to implement than those involving larger units (e.g., cities, states), given political and cost issues and methodological concerns about both the comparability of units and the ability to study a sufficient number of units to achieve adequate statistical power. That is, when aggregate units rather than individuals are randomly assigned, sample size and statistical power are based on the number of units, even if measurements are taken from individuals within each unit.

Finally, given the length of time required to plan and carry out an RCT, a trial may yield results long after the time frame in which the information is needed by decision makers. Or the results may no longer be as relevant in the context in which they would be applied when the trial has been completed. In the Women's Health Initiative Dietary Modification Trial, for example, the nature of the low-fat dietary intervention did not address important issues, related to the types of fat, that were identified while the trial was in progress and could have influenced the results (Anderson and Appel, 2006).

The above limitations can result in RCTs having less impact in research on public policy relative to other, nonrandomized designs. Consider the changes in policy that led to what is viewed as one of the major recent public health triumphs—the reduction of the amount and prevalence of tobacco smoking in the United States. None of the key large-scale interventions that were utilized—a ban on cigarette advertising, increases in the cost of tobacco products, and bans on indoor smoking in public places—were supported by direct evidence from an RCT.

Perspectives on Causal Inference

Because RCTs frequently are not possible for multicomponent interventions that deal with determinants at multiple levels (e.g., individual, family, health plan), evidence should be sought (or generated) from other designs, and some of the advantages of RCTs must be sacrificed. For example, randomized encouragement designs, in which individuals are randomized to be encouraged to take up an intervention, may allow for recruitment of a more representative sample since participants are involved in the choice of their treatment, but the level of certainty may be reduced. Cluster randomization trials may avoid the difficulties of recruiting individuals in complex settings (e.g., schoolrooms, workplaces, medical clinics, or even whole communities) while retaining randomized assignment of an intervention, but the cost or feasibility of identifying a sufficient number of groups can be prohibitive. Nonrandomized designs, with or without controls (quasi-experimental designs), may be the only feasible approach in situations where randomization is difficult or impossible. These include pre–post designs, interrupted time series, and regression discontinuity designs (elaborated in Appendix E). Such designs, when implemented well, can produce valuable evidence for effectiveness but require trade-offs with the level of certainty of causality and control of bias.

Two perspectives provide useful, complementary approaches on which researchers can draw to strengthen causal inferences, including approaches that do not involve experimentation. In the behavioral sciences, Campbell and colleagues (Campbell, 1957; Campbell and Stanley, 1966; Cook and Campbell, 1979; Shadish and Cook, 2009; Shadish et al., 2002) have developed a practical theory providing guidelines for ruling out confounders that may yield alternative explanations of research results. In the field of statistics, Rubin and colleagues (Holland, 1986; Rubin, 1974, 1978, 2005, 2008) have developed the potential outcomes perspective, which provides a deductive, mathematical approach based on making explicit, ideally verifiable assumptions. Shadish (2010), West and colleagues (2000), and West and Thoemmes (2010) offer full discussions and comparisons of these two perspectives.

Campbell's Perspective

Campbell and colleagues have considered the full range of experimental, quasi-experimental, and pre-experimental designs used by researchers in the behavioral sciences. They have identified an extensive list of threats to validity, which represent an accumulation of the various criticisms that have been made of research designs in the field of social sciences and are applicable to other fields as well (Campbell, 1988). Although Cook and Campbell (1979) describe four general types of threats to validity, the focus here is on potential confounders that undermine the level of certainty of a causal link between an intervention and the observed outcomes (internal validity) and later in the chapter is on threats that may limit the generalizability of the results (external validity). With respect to the level of certainty of causal inference, Campbell and colleagues emphasize that researchers need consider only threats to validity that are plausible given their specific design and the prior empirical research in their particular research context (Campbell, 1957; Shadish et al., 2002). Table 8-2 presents common threats to the level of certainty of causal inference associated with several major quantitative designs.

After researchers have identified the plausible threats to the level of certainty in their research context, several approaches may be taken to rule out each identified threat. First, features may be added to the design to prevent the threat from occurring. For example, a key threat in many designs is participant attrition. Shadish and colleagues (2002) describe the importance of retaining participants in a study and point to extensive protocols developed to maximize retention (see Ribisl et al., 1996). Second, certain elements may be added to the basic design. Appendix E presents several examples in which elements are added to a variety of basic designs to address those specific threats to the level of certainty that are plausible in the context of the design and prior research in the area. Shadish and Cook (1999) offer an extensive list of elements that can be included in a wide variety of research designs (see Box 8-4).

To illustrate, in the pre-experimental design with only a pretest and posttest, discussed in Appendix E, one plausible threat to a study's level of certainty is history: another event in addition to the treatment might have occurred between the pre- and posttests. Consider a year-long school-based intervention program that demonstrates a decrease in teenagers' smoking from the beginning to the end of the school year. Suppose that during the same year and unrelated to the program, the community also removed all cigarette machines that allowed children to purchase cigarettes easily. Adding the design element of replicating the study at different times in different participant cohorts would help rule out the possibility that the removal of cigarette machines rather than the school-based program was responsible for the results. If the school-based program were effective, reductions in smoking would have occurred in each replication. In contrast, removal of the cigarette machines would be expected to lead to a decrease only in the first replication, a different expected pattern of results. Matching the pattern of results to that predicted by the theory of the intervention versus that predicted by the plausible confounders provides a powerful method of ruling out threats to a study's level of certainty.

Campbell's approach strongly prefers such design strategies over alternative statistical adjustment strategies to deal with threats to the level of certainty. It also emphasizes strategies for increasing researchers' understanding of the particular conceptual aspects of the treatment that are responsible for the causal effects under the rubric of the construct validity of the independent variable. Shadish and colleagues (2002) present a general discussion of these issues, and West and Aiken (1997) and Collins and colleagues (2009) discuss experimental designs for studying the effective-

TABLE 8-2 Key Assumptions/Threats to Internal Validity[a] and Example Remedies for Randomized Controlled Trials and Alternatives

Assumption or Threat to Internal Validity	Approaches to Mitigating the Threat to Internal Validity: Design Approach	Statistical Approach
Randomized controlled experiment		
Independent units	Temporal or geographical isolation of units	Multilevel analysis (other statistical adjustment for clustering)
Full treatment adherence	Incentives for adherence	Instrumental variable analysis (assume exclusion restriction)
No attrition	Sample retention procedures	Missing data analysis (assume missing at random)
Other treatment conditions do not affect participant's outcome (SUTVA)	Temporal or geographical isolation of treatment groups	Statistical adjustment for measured exposure to other treatments
Randomized encouragement design		
Exclusion restriction	No design approach yet available	Sensitivity analysis
Regression discontinuity design		
Functional form of the relationship between assignment variable and outcome is properly specified	Replication with different threshold; nonequivalent dependent variable	Nonparametric regression; sensitivity analysis
Interrupted time series analysis		
Functional form of the relationship for the time series is properly specified; another historical event, a change in population (selection), or a change in measures coincides with the introduction of the intervention	Nonequivalent control series in which intervention is not introduced; switching replication in which intervention is introduced at another time point; nonequivalent dependent measure	Diagnostic plots (autocorrelogram; spectral density); sensitivity analysis
Observational study		
Measured baseline variables equated; unmeasured baseline variables equated; differential maturation; baseline variables reliably measured	Multiple control groups; nonequivalent dependent measures; additional pre- and postintervention measurements	Propensity score analysis; sensitivity analysis; subgroup analysis; correction for measurement error

[a] Internal validity is defined as a study's level of certainty of the causal relationship between an intervention and the observed outcomes.
NOTE: SUTVA = stable unit treatment value assumption. The list of assumptions and threats to internal validity identifies issues that commonly occur in each of the designs. The alternative designs may be subject to each of the issues listed for the randomized conrolled trial in addition to the issues listed for the specific design. The examples of statistical and design approaches for mitigating the threat to internal validity illustrate some commonly used approaches and are not exhaustive. For the observational study design, the potential outcomes and Campbellian frameworks study differ so that the statistical and design approaches do not map 1-to-1 onto the assumptions or threats to internal validity that are listed.
SOURCE: West et al., 2008. Reprinted with permission. West et al., Alternatives to the randomized controlled trial, *American Journal of Public Health*, 98(8):1364, Copyright © 2008 by the American Public Health Association.

ness of individual components in multiple-component designs. West and Aiken (1997) and MacKinnon (2008) also consider the statistical method of mediational analysis that permits researchers to probe these issues, although in a less definitive manner.

Rubin's Perspective

Rubin's potential outcomes model takes a formal mathematical/statistical approach to causal inference. Building on earlier work by Splawa-Neyman (1990), it emphasizes precise definition of the desired causal effect and specification of explicit, ideally verifiable assumptions that are sufficient to draw causal inferences for each research design. Rubin defines a causal effect as the difference between the outcomes for a single unit (e.g., person, community) given two different well-defined treatments at the identical time and in the same context. This definition represents a conceptually useful ideal that cannot be realized in practice.

Holland (1986) notes that three approaches, each with its own assumptions, can be taken to approximate this ideal. First, a within-subjects design can be used in which the two treatments (e.g., intervention, control) are given to the same unit. This design assumes (1) temporal stability in which the same outcome will be observed regardless of when the treatment is delivered and (2) causal transience in which the administration of the first treatment has no effect on the outcome of the second treatment. These assumptions will frequently be violated in research on obesity. Second, homogeneous units can be selected or created so that each unit can be expected to have the same response to the treatment. This strategy is commonly used in engineering applications, but raises concern about the comparability of units in human research—even monozygotic twins raised in similar environments can differ in important ways in some research contexts. The matching procedures used in the potential outcomes approach discussed below rely on this approach; they assume that the units can indeed be made homogeneous on all potentially important background variables. Third, units can be randomly assigned to treatment and control conditions. This strategy creates groups that are, on average, equal on all possible background variables at pretest so that the difference between the means of the two groups now represents the *average causal effect*. This strategy makes several assumptions (see Table 8-2; Holland, 1986; West and Thoemmes, 2010), including full treatment adherence, independence of units, no attrition from posttest measurement, and the nondependence of the response of a unit to a treatment on the treatment received by other units (or SUTVA, the stable unit treatment value assumption). The SUTVA highlights the challenges in community research of considering possible dependence between units and possible variation in each treatment across sites (Rubin, 2010). Well-defined treatment (and no-treatment) conditions that are implemented identically across units are a key feature of strong causal inference in Rubin's perspective. Hernan and Traubman (2008) discuss the importance of this assumption in the context of obesity research. Beyond requiring this set of foundational assumptions, randomization has another subtle effect: it shifts the focus from a causal effect defined at the level of the individual to an

Box 8-4
Design Elements Used in Constructing Quasi-Experiments

Assignment (control of assignment strategies to increase group comparability)

- Cutoff-based assignment: controlled assignment to conditions based solely on one or more fully measured covariates; this yields an unbiased effect estimate.
- Other nonrandom assignment: various forms of "haphazard" assignment that sometimes approximate randomization (e.g., alternating assignment in a two-condition quasi-experiment whereby every other unit is assigned to one condition).

Measurement (use of measures to learn whether threats to causal inference actually operate)

- Posttest observations
  - Nonequivalent dependent variables: measures that are not sensitive to the causal forces of the treatment, but are sensitive to all or most of the confounding causal forces that might lead to false conclusions about treatment effects (if such measures show no effect, but the outcome measures do show an effect, the causal inference is bolstered because it is less likely due to the confounds).
  - Multiple substantive posttests: used to assess whether the treatment affects a complex pattern of theoretical predicted outcomes.
- Pretest observations
  - Single pretest: a pretreatment measure on the outcome variable, useful to help diagnose selection bias.
  - Retrospective pretest: reconstructed pretests when actual pretests are not feasible—by itself, a very weak design feature, but sometimes better than nothing.
  - Proxy pretest: when a true pretest is not feasible, a pretest on a variable correlated with the outcome—also often weak by itself.
  - Multiple pretest time points on the outcome: helps reveal pretreatment trends or regression artifacts that might complicate causal inference.
- Pretests on independent samples: when a pretest is not feasible on the treated sample, one is obtained from a randomly equivalent sample.
- Complex predictions such as predicted interaction: successfully predicted interactions lend support to causal inference because alternative explanations become less plausible.

average causal effect that characterizes the difference between the treatment and control groups. The ability to make statements about individual causal effects, important in many clinical and health contexts, is diminished without additional assumptions being made (e.g., the causal effect is constant for all individuals).

A key idea in Rubin's perspective is that of possible outcomes. The outcome of a single unit (participant) receiving a treatment is compared with the outcome that would have occurred if the same unit had received the alternative treatment. This idea has proven to be a remarkably generative way of thinking about how to obtain precise estimates of causal effects. It focuses the researcher on the precise nature of the comparison that needs to be made and clearly delineates the participants for whom the comparison is appropriate. It provides a basis for elegant solutions to such problems as treatment nonadherence and appropriate matching in nonrandomized studies (West and Thoemmes, 2010).

- Measurement of threats to internal validity:[a] helps diagnose the presence of specific threats to the inference that A caused B, such as whether units actively sought out additional treatments outside the experiment.

Comparison groups (selecting comparisons that are "less nonequivalent" or that bracket the treatment group at the pretest[s])

- Single nonequivalent groups: compared with studies without control groups, using a nonequivalent control group helps identify many plausible threats to validity.
- Multiple nonequivalent groups: serve several functions; for instance, groups are selected that are as similar as possible to the treated group, but at least one outperforms it initially and at least one underperforms it, thus bracketing the treated group.
- Cohorts: comparison groups chosen from the same institute in a different cycle (e.g., sibling controls in families or last year's students in schools).
- Internal (versus external) controls: plausibly chosen from within the same population (e.g., within the same school rather than from a different school).

Treatment (manipulations of the treatment to demonstrate that treatment variability affects outcome variability)

- Removed treatments: shows that an effect diminishes if treatment is removed.
- Repeated treatments: reintroduces treatments after they have been removed from some group—common in laboratory science.
- Switched replications: treatment and control group roles are reversed so that one group is the control while the other receives treatment, but the controls receive treatment later, whereas the original treatment group receives no further treatment or has treatment removed.
- Reversed treatments: provides a conceptually similar treatment that reverses an effect—for example, reducing access for some students to a computer being studied by increasing access for others.
- Dosage variation (treatment partitioning): demonstrates that an outcome responds systematically to different levels of treatment.

[a] Internal validity denotes the level of certainty of the causal relationship between an intervention and the observed outcomes.

SOURCE: Reprinted, with permission. Shadish and Cook, 1999. Copyright © 1999 by the Institute of Mathematical Statistics.

Other Perspectives on Causality

Other perspectives on causal inference exist. In economics, Granger (1988; Granger and Newbold, 1977; see also Bollen, 1989, in sociology and Kenny, 1979, in psychology) argues that three conditions are necessary to infer that one variable X causes changes in another variable Y:

- Association—The two variables X and Y must be associated (nonlinear association is permitted).
- Temporal precedence—X must precede Y in time.
- Nonspuriousness—X contains unique information about Y that is not available elsewhere. Otherwise stated, with all other causes partialed out, X still predicts Y.

This perspective can be particularly useful for thinking about causal inference when X cannot be manipulated (e.g., age), but is typically less useful for studying the effects of interventions. The primary challenge is how one can establish nonspuriousness, and Granger's perspective provides few guidelines in this regard relative to the Campbell and Rubin perspectives.

A second class of perspectives, founded in philosophical and computer science, takes a graph theoretic approach. In this approach, a complex model of the process is specified and is compared with data. If the model and its underlying assumptions are true, the approach can discern whether causal inferences can be supported. Within this approach, a computer program known as Tetrad (Spirtes et al., 2000) can identify any other models involving the set of variables in the system that provide an equally good account of the data, if they exist. In separate work, Pearl (2009) has also utilized the graph theoretic approach, developing a mathematical calculus for problems of causal inference. This approach offers great promise as a source of understanding of causal effects in complex systems in the future. Compared with the Campbell and Rubin approaches, however, to date it has provided little practical guidance for researchers attempting to strengthen the inferences about the effectiveness of interventions that can be drawn from evaluation studies.

How Well Do Alternative Designs Work?

Early attempts to compare the magnitude of the causal effects estimated from RCTs and nonrandomized designs used one of two approaches. First, the results from an RCT and a separate observational study investigating the same question were compared. For example, LaLonde (1986) found that an RCT and nonrandomized evaluations of the effectiveness of job training programs led to completely different results. Second, the results of interventions in a research area evaluated using randomized and nonrandomized designs were compared in a meta-analytic review. For example, Sacks and colleagues (1983) compared results of RCTs of medical interventions with results of nonrandomized designs using historical controls, finding that the nonrandomized designs overestimated the effectiveness of the interventions. A number of studies showing the noncomparability of results of RCTs and nonrandomized designs exist, although many of the larger meta-analyses in medicine (Ioannidis et al., 2001) and the behavioral sciences (Lipsey and Wilson, 1993) find no evidence of consistent bias.

More recently, Cook and colleagues (2008) compared the small set of randomized and nonrandomized studies that shared the same treatment group and the same measurement of the outcome variable. All cases in which an RCT was compared with a regression discontinuity or interrupted time series study design (see Appendix E for discussion of these study designs) showed no differences in effect size. Observational studies produced more variable results, but the results did not differ from those of an RCT given that (1) a control group of similar participants was employed or (2) the mechanism for selection into treatment and control groups was known. Hernán and colleagues (2008) considered the disparate results of an RCT (Women's Health Initiative) and an observational study (Nurses' Health Study) of the effectiveness of postmenopausal hormone replacement therapy. When the observational study was analyzed using propensity score methods with participants who met the same eligibility criteria and the same intention-to-treat causal effect was estimated, discrepancies were minimal. Finally, Shadish and colleagues (2008) randomly assigned student participants to an RCT or observational study (self-selection) of the effects of math or vocabulary training. They found little difference in the estimates of causal effects after adjusting for an extensive set of baseline covariates in the observational study.

The results of the small number of focused comparisons of randomized and nonrandomized designs to date are encouraging. Additional research comparing treatment effect estimates for randomized and nonrandomized designs using similar treatments, populations of participants, and effect size estimates is needed to determine the generality of this finding.

REFERENCES

Anderson, C. A., and L. J. Appel. 2006. Dietary modification and CVD prevention: A matter of fat. *Journal of the American Medical Association* 295(6):693-695.

Bollen, K. A. 1989. *Structural equations with latent variables*. New York: John Wiley and Sons, Inc.

Bonell, C., J. Hargreaves, V. Strange, P. Pronyk, and J. Porter. 2006. Should structural interventions be evaluated using RCTs? The case of HIV prevention. *Social Science and Medicine* 63(5):1135-1142.

Campbell, D. T. 1957. Factors relevant to the validity of experiments in social settings. *Psychological Bulletin* 54(4):297-312.

Campbell, D. T. 1988. Can we be scientific in applied social science? In *Methodology and epistemology for social science: Selected papers of Donald T. Campbell*, edited by E. S. Overman. Chicago, IL: University of Chicago Press. Pp. 315-334.

Campbell, D. T., and J. C. Stanley. 1966. *Experimental and quasi-experimental designs for research*. Chicago, IL: Rand McNally.

CBO (Congressional Budget Office). 2008. *Key issues in analyzing major health insurance proposals*. Publication No. 3102. Washington, DC: U.S. Congress.

CDC (Centers for Disease Control and Prevention). 1999. *Best practices for comprehensive tobacco control programs—April 1999*. Atlanta, GA: U.S. Department of Health and Human Services, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health.

Collins, L. M., J. J. Dziak, and R. Li. 2009. Design of experiments with multiple independent variables: A resource management perspective on complete and reduced factorial designs. *Psychological Methods* 14(3):202-224.

Cook, T. D., and D. T. Campbell. 1979. *Quasi-experimentation: Design and analysis issues for field settings*. Chicago, IL: Rand McNally College Publishing Co.

Cook, T. D., W. R. Shadish, and V. C. Wong. 2008. Three conditions under which experiments and observational studies produce comparable causal estimates: New findings from within-study comparisons. *Journal of Policy Analysis and Management* 27(4):724-750.

Donner, A., and N. Klar. 2000. *Design and analysis of cluster randomization trials in health research*. London: Arnold.

Edwards, N. M., and S. J. Schwarzenberg. 2009. Designing and implementing an effective pediatric weight management program. *Reviews in Endocrine and Metabolic Disorders* 10(3):197-203.

Eriksen, M. 2006. Are there public health lessons that can be used to help prevent childhood obesity? *Health Education Research* 21(6):753-754.

Garfield, S. A., S. Malozowski, M. H. Chin, K. M. Narayan, R. E. Glasgow, L. W. Green, R. G. Hiss, and H. M. Krumholz. 2003. Considerations for diabetes translational research in real-world settings. *Diabetes Care* 26(9):2670-2674.

Glasgow, R. E., E. Lichtenstein, and A. C. Marcus. 2003. Why don't we see more translation of health promotion research to practice? Rethinking the efficacy-to-effectiveness transition. *American Journal of Public Health* 93(8):1261-1267.

Glasgow, R., L. Green, L. Klesges, D. Abrams, E. Fisher, M. Goldstein, L. Hayman, J. Ockene, and C. Olrleans. 2006. External validity: We need to do more. *Annals of Behavioral Medicine* 31(2):105-108.

Glasgow, R. E., L. W. Green, and A. Ammerman. 2007. A focus on external validity. *Evaluation and the Health Professions* 30(2):115-117.

Granger, C. W. J. 1988. Some recent development in a concept of causality. *Journal of Econometrics* 39(1-2):199-211.

Granger, C. W. J., and P. Newbold. 1977. *Forecasting economic time series*. New York: Academic Press.

Green, L. W. 2001. From research to "best practices" in other settings and populations. *American Journal of Health Behavior* 25(3):165-178.

Green, L. W. 2007. The Prevention Research Centers as models of practice-based evidence: Two decades on. *American Journal of Preventive Medicine* 33(1, Supplement):S6-S8.

Green, L. W., and R. E. Glasgow. 2006. Evaluating the relevance, generalization, and applicability of research: Issues in external validation and translation methodology. *Evaluation & the Health Professions* 29(1):126-153.

Green, L. W., and F. M. Lewis. 1981. Issues in relating evaluation to theory, policy, and practice in continuing education and health education. *Mobius* 81(1):46-58.

Green, L. W., F. M. Lewis, and D. M. Levine. 1980. Balancing statistical data and clinician judgments in the diagnosis of patient educational needs. *Journal of Community Health* 6(2):79-91.

Green, L. W., R. E. Glasgow, D. Atkins, and K. Stange. 2009. Making evidence from research more relevant, useful, and actionable in policy, program planning, and practice: Slips "twixt cup and lip". *American Journal of Preventive Medicine* 37(6, Supplement 1):S187-S191.

Hay, M. C., T. S. Weisner, S. Subramanian, N. Duan, E. J. Niedzinski, and R. L. Kravitz. 2008. Harnessing experience: Exploring the gap between evidence-based medicine and clinical practice. *Journal of Evaluation in Clinical Practice* 14(5):707-713.

Hernan, M. A., and S. L. Taubman. 2008. Does obesity shorten life? The importance of well-defined interventions to answer causal questions. *International Journal of Obesity (London)* 32(Supplement 3):S8-S14.

Hernan, M. A., A. Alonso, R. Logan, F. Grodstein, K. B. Michels, W. C. Willett, J. E. Manson, and J. M. Robins. 2008. Observational studies analyzed like randomized experiments: An application to postmenopausal hormone therapy and coronary heart disease. *Epidemiology* 19(6):766-779.

Holland, P. W. 1986. Statistics and causal inference (with discussion). *Journal of the American Statistical Association* 81(396):945-960.

Ioannidis, J. P. A., A. B. Haidich, M. Pappa, N. Pantazis, S. I. Kokori, M. G. Tektonidou, D. G. Contopoulos-Ioannidis, and J. Lau. 2001. Comparison of evidence of treatment effects in randomized and nonrandomized studies. *Journal of the American Medical Association* 286(7):821-830.

IOM (Institute of Medicine). 2007. *Progress in preventing childhood obesity: How do we measure up?* Edited by J. Koplan, C. T. Liverman, V. I. Kraak, and S. L. Wisham. Washington, DC: The National Academies Press.

Katz, R. J. 2009. Communities and academia working together: Report of the Association of Schools of Public Health Prevention Research Centers Blue Ribbon Panel. *Public Health Reports* 124(2):334-338.

Kenny, D. A. 1979. *Correlation and causality.* New York: John Wiley Interscience.

Kessel, F., and P. L. Rosenfield. 2003. Fostering interdisciplinary innovation: The way forward. In *Expanding the boundaries of health and social science: Case studies of interdisciplinary innovation*, edited by F. Kessel, P. L. Rosenfield and N. B. Anderson. New York: Oxford University Press, Inc. p. 401.

Kessel, F., and P. L. Rosenfield. 2008. Toward transdisciplinary research: Historical and contemporary perspectives. *American Journal of Preventive Medicine* 35(2 Supplement): S225-S234.

Klesges, L. M., D. A. Dzewaltowski, and R. E. Glasgow. 2008. Review of external validity reporting in childhood obesity prevention research. *American Journal of Preventive Medicine* 34(3):216-223.

Kottke, T. E., L. I. Solberg, A. F. Nelson, D. W. Belcher, W. Caplan, L. W. Green, E. Lydick, D. J. Magid, S. J. Rolnick, and S. H. Woolf. 2008. Optimizing practice through research: A new perspective to solve an old problem. *Annals of Family Medicine* 6(5):459-462.

Kriemler, S., L. Zahner, C. Schindler, U. Meyer, T. Hartmann, H. Hebestreit, H. P. Brunner-La Rocca, W. van Mechelen, and J. J. Puder. 2010. Effect of school based physical activity programme (KISS) on fitness and adiposity in primary schoolchildren: Cluster randomised controlled trial. *British Medical Journal* 340:c785.

LaLonde, R. J. 1986. Evaluating the econometric evaluations of training programs with experimental data. *American Economic Review* 76(4):604-620.

Leviton, L. 2001. External validity. In *International encyclopedia of the behavioral and social sciences*. Vol. 8, edited by N. J. Smelser and P. B. Baltes. Oxford: Elsevier. Pp. 5195-5200.

Leviton, L. C., C. B. Collins, B. L. Liard, and P. P. Kratt. 1998. Teaching evaluation using evaluability assessment. *Evaluation* 4(4):389-409.

Leviton, L., L. Kettel Khan, D. Rog, N. Dawkins, and D. Cotton. 2010a. Exploratory evaluation of public health policies, programs, and practices. *Annual Review of Public Health* 31:213-233.

Leviton, L. C., L. Kettel Khan, and N. Dawkins (editors). 2010b. The systematic screening and assessment method: Finding innovations worth evaluating. *New directions in evaluation* 2010(125):1-118.

Lipsey, M. W., and D. B. Wilson. 1993. The efficacy of psychological, educational, and behavioral treatment: Confirmation from meta-analysis. *American Psychologist* 48(12):1181-1209.

MacKinnon, D. P. 2008. *An introduction to statistical meditation analysis*. New York: Lawrence Erlbaum Associates.

Mercer, S. L., B. J. DeVinney, L. J. Fine, L. W. Green, and D. Dougherty. 2007. Study designs for effectiveness and translation research: Identifying trade-offs. *American Journal of Preventive Medicine* 33(2):139-154.

Mullen, P. D., L. W. Green, and G. S. Persinger. 1985. Clinical trials of patient education for chronic conditions: A comparative meta-analysis of intervention types. *Preventive Medicine* 14(6):753-781.

Murray, D. M. 1998. *Design and analysis of group randomized trials*. New York: Oxford.

Ockene, J. K., E. A. Edgerton, S. M. Teutsch, L. N. Marion, T. Miller, J. L. Genevro, C. J. Loveland-Cherry, J. E. Fielding, and P. A. Briss. 2007. Integrating evidence-based clinical and community strategies to improve health. *American Journal of Preventive Medicine* 32(3):244-252.

Patrick, K., F. D. Scutchfield, and S. H. Woolf. 2008. External validity reporting in prevention research. *American Journal of Preventive Medicine* 34(3):260-262.

Pearl, J. 2009. *Causality: Models, reasoning, and interference*. 2nd ed. New York: Cambridge.

Pratt, C. A., J. Stevens, and S. Daniels. 2008. Childhood obesity prevention and treatment: Recommendations for future research. *American Journal of Preventive Medicine* 35(3):249-252.

Ramanathan, S., K. R. Allison, G. Faulkner, and J. J. Dwyer. 2008. Challenges in assessing the implementation and effectiveness of physical activity and nutrition policy interventions as natural experiments. *Health Promotion International* 23(3):290-297.

Raudenbush, S. W., and X. Liu. 2000. Statistical power and optimal design for multisite randomized trials. *Psychological Methods* 5(2):199-213.

Ribisl, K. M., M. A. Walton, C. T. Mowbray, D. A. Luke, W. S. Davidson II, and B. J. Bootsmiller. 1996. Minimizing participant attrition in panel studies through the use of effective retention and tracking strategies: Review and recommendations. *Evaluation and Program Planning* 19(1):1-25.

Rosenfield, P. L. 1992. The potential of transdisciplinary research for sustaining and extending linkages between the health and social sciences. *Social Science and Medicine* 35(11):1343-1357.

Rubin, D. B. 1974. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology* 66(5):688-701.

Rubin, D. B. 1978. Bayesian inference for causal effects: The role of randomization. *Annals of Statistics* 6(1):34-58.

Rubin, D. B. 2005. Causal inference using potential outcomes: Design, modeling, decisions. *Journal of the American Statistical Association* 100(469):322-331.

Rubin, D. B. 2008. Statistical inference for causal effects, with emphasis on applications in epidemiology and medical statistics. In *Handbook of statistics 27: Epidemiology and medi-*

cal statistics, edited by C. R. Rao, J. P. Miller and D. C. Rao. Amsterdam, the Netherlands: Elsevier. Pp. 28-63.

Rubin, D. B. 2010. Reflections stimulated by the comments of Shadish (2010) and West and Thoemmes (2010). *Psychological Methods* 15(1):38-46.

RWJF (Robert Wood Johnson Foundation). 2006. *Early assessment of environmental interventions to prevent childhood obesity. meeting summary, January 19–20, 2006*. http://www.rwjf.org/files/publications/other/EA%20Report%200612.pdf (accessed January 4, 2010).

RWJF. 2009. *Early assessment of programs and policies on childhood obesity*. http://www.rwjf.org/pr/product.jsp?id=37330 (accessed January 4, 2010).

Sacks, H. S., T. C. Chalmers, and H. Smith Jr. 1983. Sensitivity and specificity of clinical trials: Randomized v historical controls. *Archives of Internal Medicine* 143(4):753-755.

Shadish, W. R. 2010. Campbell and Rubin: A primer and comparison of their approaches to causal inference in field settings. *Psychological Methods* 15(1):3-17.

Shadish, W. R., and T. D. Cook. 1999. Design rules: More steps toward a complete theory of quasi-experimentation. *Statistical Science* 14:294-300.

Shadish, W. R., and T. D. Cook. 2009. The renaissance of field experimentation in evaluating interventions. *Annual Review of Psychology* 60:607-629.

Shadish, W. R., T. D. Cook, and D. T. Campbell. 2002. *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin.

Shadish, W. R., M. H. Clark, and P. M. Streiner. 2008. Can nonrandomized experiments yield accurate answers? A randomized experiment comparing random and nonrandom assignments. *Journal of the American Statistical Association* 103:1334-1356.

Shepherd, W. L., and A. Moore. 2008. Task force recommendations: Application in the "real world" of community intervention. *American Journal of Preventive Medicine* 35(1, Supplement):S1-S2.

Simons-Morton, D. G., P. D. Mullen, D. A. Mains, E. R. Tabak, and L. W. Green. 1992. Characteristics of controlled studies of patient education and counseling for preventive health behaviors. *Patient Education and Counseling* 19(2):175-204.

Spirtes, P., C. Glymour, and R. Scheines. 2000. *Causation, prediction, and search*. 2nd ed. New York: MIT Press.

Splawa-Neyman, J., D. M. Dabrowska, and T. P. Speed. 1990. On the application of probability theory to agricultural experiments. Essay on principles. Section 9. *Statistical Science* 5(4):465-472.

Steckler, A., and K. R. McLeroy. 2008. The importance of external validity. *American Journal of Public Health* 98(1):9-10.

Tabak, E. R., P. D. Mullen, D. G. Simons-Morton, L. W. Green, D. A. Mains, S. Eilat-Greenberg, R. F. Frankowski, and M. C. Glenday. 1991. Definition and yield of inclusion criteria for a meta-analysis of patient education studies in clinical preventive services. *Evaluation and the Health Professions* 14(4):388-411.

Varnell, S. P., D. M. Murray, J. B. Janega, and J. L. Blitstein. 2004. Design and analysis of group-randomized trials: A review of recent practices. *American Journal of Public Health* 94(3):393-399.

Weisz, J. R., B. Weiss, and G. R. Donenberg. 1992. The lab versus the clinic: Effects of child and adolescent psychotherapy. *American Psychologist* 47(12):1578-1585.

West, S. G., and L. S. Aiken. 1997. Towards understanding individual effects in multiple component prevention programs: Design and analysis strategies. In *Recent advances in prevention methodology: Alcohol and substance abuse research*, edited by K. Bryant, M. Windle, and S. G. West. Washington, DC: American Psychological Association.

West, S. G., and F. Thoemmes. 2010. Campbell's and Rubin's perspectives on causal inference. *Psychological Methods* 15(1):18-37.

West, S., J. Biesanz, and S. Pitts. 2000. Causal inference and generalization in field settings: Experimental and quasi-experimental designs. In *Handbook of research methods in social and personality psychology*, edited by H. Reis and C. Judd. Cambridge: University Press. Pp. 40-84.

West, S., J. Duan, W. Pequegnat, P. Gaist, D. Des Jarlais, D. Holtgrave, J. Szapocznik, M. Fishbein, B. Rapkin, M. Clatts, and P. Mullen. 2008. Alternatives to the randomized controlled trial. *American Journal of Public Health* 98(8):1359-1366.

Wholey, J. 1977. Evaluability assessment. In *Evaluation research methods: A basic guide*, edited by L. Rutman. Beverly Hills, CA: Sage. Pp. 41-56.

Wholey, J. 1994. Assessing the feasibility and likely usefulness of evaluation. In *The handbook of practical program evaluation*, edited by J. S. Wholey, H. P. Hatry, and K. E. Newcomer. San Francisco, CA: Jossey-Bass.

Yanek, L. R., D. M. Becker, T. F. Moy, J. Gittelsohn, and D. M. Koffman. 2001. Project joy: Faith-based cardiovascular health promotion for African American women. *Public Health Reports* 116(Supplement 1):68-81.

9

Next Steps

KEY MESSAGES

Communication and Dissemination Plan

- Decision makers (e.g., policy makers and their intermediaries, program planners, and practitioners) and researchers working on obesity prevention and other complex public health problems are the target audiences for communicating the L.E.A.D. framework. Other important audiences for the framework include research funders and publishers.
- It is important to understand the settings, communication channels, and activities of these audiences to engage and educate them effectively on the purpose and adoption of the L.E.A.D. framework.
- To support the development and refinement of a communication and dissemination plan, it will be necessary to create partnerships, use existing activities and networks, and tailor the messages and approaches to each target audience.

Evaluation and Refinement Plan

- Key outcome measures—utilization, adoption, acceptance, maintenance, and impact—should be defined and evaluated to assess the integration of the L.E.A.D. framework into the thinking and practice of the target audiences.
- Improvement or refinement of the L.E.A.D. framework and its further integration into practice will require making use of data collection methods and opportunities that will best suit the target audiences. Evaluation of the effects of using the framework should also be supported so that feedback from actual decision-making settings can be integrated.

The committee's statement of task included developing a plan for communicating and disseminating the framework proposed in this report and specifying a plan for evaluating and refining the framework in the context of current decision-making processes (see Chapter 1). As a result of its deliberations, the committee determined that the plans for communication and dissemination and for evaluation and refinement specified in its statement of task would best be developed in the context of other aspects of the framework's practical application as defined by its potential users—decision makers and researchers—who will identify, evaluate, generate, incorporate, or use evidence to inform their selection, implementation, and evaluation of obesity prevention efforts. To guide future efforts, this chapter outlines key considerations for the development of these plans.

COMMUNICATING AND DISSEMINATING THE FRAMEWORK

Objectives

If the L.E.A.D. framework is to be communicated and disseminated successfully to its target audiences, three major objectives need to be met. First, an environment has to be created in which the framework and its elements are known and understood. Second, a culture of shared beliefs about what constitutes evidence and how it ought to be used to make decisions should be created in accordance with the framework. Finally, to develop support for implementation of the framework, strong champions need to be recruited to lay the foundation for its integration into public health decision making and related research. To develop a communication and dissemination plan that meets these objectives, it will be necessary to define and learn about the target audiences, explore the channels that are best suited to reaching those audiences, identify potential partnerships, determine the messages to be delivered and the materials that will best convey them, and decide how and when to implement the plan and how to monitor its implementation (NCI, 2008).

Target Audiences

The target audiences for the L.E.A.D. framework are decision makers (e.g., policy makers and their intermediaries, program planners, and practitioners) and researchers working on obesity prevention and other complex public health problems. Research funders and publishers of research results are also important audiences for the framework.

Decision Makers and Intermediaries

A variety of interested individuals and organizations help identify public health issues and potential strategies for addressing these issues through policy decisions. They include public health departments; local, state, and federal policy makers; health plans; employers; and professional and trade organizations. They also include the staff of decision makers, consultants, and firms that specialize in gathering and synthesizing research, as well as academic and private foundation policy think tanks and advocacy and stakeholder groups. These intermediaries may be involved in or belong to professional organizations and associations and network with those who have similar professional interests (e.g., employee business groups; state, local, or community legislation–based groups). They vary considerably in how, why, and when they use and incorporate evidence to help them develop strategies for addressing a particular problem.

Researchers

Evidence is generated by individuals and organizations that conduct research. Researchers do their work in both the public and private sectors, including agencies within local and federal government, academia, foundations, and industry. Many are affiliated with the setting where they perform the research and use their work-affiliated networks, as well as professional organizations and associations to which they belong, to obtain and circulate information. The research they generate is influenced by the resources available to them or the resources they themselves can garner to support their work.

Other Important Audiences

Research Funders. Research is supported by funders that supply the necessary equipment, money, facilities, tools, and other materials. Funders of research also exist throughout the public and private sectors. The main supporters of research in the public sector include local and federal government agencies that focus on public health issues. These agencies, as well as large organizations in the private sector, support the majority of public health–related research. Their decisions to support research on particular topics or groups are based on a variety of protocols and policies their leaders have developed to meet their organizations' objectives. State or federal legislatures may mandate that certain agencies produce reports on topics of particular importance to their constituencies and may have specific restrictions on how and for what purpose the grants they offer can be used. However, many of the factors that affect decisions to support research are not apparent. For example, private organizations may have policy restrictions on the research they will support, but also have the flexibility to change the focus of their support to meet the needs of their targeted communities or staff members.

Many federal, state, and local government staff members belong to and attend the meetings of a number of field-specific professional organizations. They may also network among themselves through interagency task forces and with the rest of the community through regular meetings with their grantees. A number of private organizations and industry groups are members of or support coalitions and associations as well. For example, the Healthy Eating Active Living Convergence Partnership (HEALCP) comprises a number of private organizations and a federal agency that have come together "with the shared goal of changing policies and environments to better achieve the vision of healthy people living in healthy places" (HEALCP, 2008).

Publishers of Research Results. Once the research is complete, its results are published by various public and private entities. Publishers of research results play an important role in determining what information becomes available; without a venue for publication, researchers' findings and conclusions may not reach the public. A number of public and private organizations and academia help distribute the results of research studies, including the funders of the research or those with an interest or specialization in the topic. Peer-reviewed journals, professional organizations (and their associated publications), foundations, media, industry, and the public sector (local, state, and federal government) all publish research results. Some apply a strict set of policies to establish standards for the research they publish, while others may not have such policies in place. Additionally, publishers establish the accessibility of the research they publish. They determine whether it is free to the public, whether it costs a nominal amount, or whether only their associated members can access it. Before the advent of the Internet and powerful search engines, research reached the public primarily through the medium of scientific journals. Now, web-based venues are accessible to search engines, and researchers have several channels through which to communicate their findings. However, peer-reviewed journals have much more credibility, at least to scientists and decision makers, than other Internet sources.

Channels

A variety of channels can be used for communicating and disseminating the L.E.A.D. framework, each of which has both benefits and drawbacks. For example, websites can reach a large number of people but require maintenance over time.

Individualized approaches may be necessary to help meet the objectives of the communication and dissemination plan. Although the time and resources required for these approaches may be prohibitive, individualized messaging and materials are among the most effective ways to affect attitudes, skills, and behavior and are more likely to be trusted and influential than other information channels (NCI, 2008). Key researchers with a major influence on the practice and policy environment should be identified and individually targeted to gain their support for integrating the framework into their current practices. For example, informal discussions and training sessions could be held for these key audiences.

At the same time, organizational-level approaches will be required to help meet the objectives of the framework's communication and dissemination strategy. Established meetings and conferences are a familiar and trusted source of information, training, and best practices for many members of the target audiences for the framework. This venue can reach a large audience in one place cost-effectively. For example, each year groups such as the U.S. Centers for Disease Control and Prevention (CDC) and the National Chronic Disease Directors sponsor meetings that are attended by key personnel from state programs. These meetings often offer training sessions for public health professionals (including program and/or evaluation staff), partners, and stakeholders and would be ideal venues for communicating the framework. Additionally, research funders and decision makers may convene or sponsor conferences and workshops to help disseminate research findings and identify priority areas for future research. Conferences and workshops that are sponsored by well-respected supporters of particular fields of public health are especially well attended. These meetings may be either open to the public (usually government-related), attended by affiliated members of the sponsoring organization (trade or professional), or attended only by selected stakeholders invited by the sponsoring organization. Journal editors may attend such meetings as well to identify research results they want to publish and to contact the researchers about submitting an article.

Writing articles for professional peer-reviewed journals or newsletters is another way to reach a large proportion of the framework's target audiences. Although the writing and publication process may be time-consuming, using such trusted sources of information to convey the messages of the framework can be particularly influential. Another approach to delivering the messages of the framework is through funders of obesity prevention research, which can support the use, understanding, and evaluation of the framework by incorporating it into requests for proposals.

Although individual and organizational channels may be more effective, mass media channels should also be incorporated into the communication and dissemination strategy for the L.E.A.D. framework. Through the use of interactive media such as websites, webinars, CD-ROMs, and e-mail mailing lists, it is possible to reach a large number of people quickly. Through the development of interactive and graphically appealing materials, messages can be tailored to meet the needs of and appeal to multiple members of the framework's target audiences. On the other hand, these channels can be expensive to develop and maintain over time; effort is also required to reach an audience with these messages, instead of depending on the audience to search actively for them.

Partnerships

Working with other organizations can be a cost-effective way to extend the reach of the framework, enhance the credibility of its messages, and create support for implementing and sustaining the concepts it embodies. Through such partnering, it is possible to tap into an organization's resources, expertise, and networks. However, it is important to develop a specific plan with potential partners to ensure that all involved have the same goals and to specify the role each partner will take in communicating and disseminating the framework.

Several public–private partnerships specifically address obesity prevention–related research and activities. They include HEALCP and the National Collaborative on Childhood Obesity Research (NCCOR). Knowing the mission and objectives of such collaborations can help in identifying their potential role in the communication and dissemination of the framework. HEALCP is a collaboration of funders whose goal is to change policies and environments to achieve healthy people and healthy places, while NCCOR is focused on improving the efficiency, effectiveness, and application of obesity research (specifically for children) (HEALCP, 2008; NCCOR, 2009). Such existing collaborations are important sources of support for integrating the framework into public health practices.

Individually, public and private organizations and agencies also have existing networks and expertise that may be important sources of partners for communicating and disseminating the framework. For example, the scope of work in a number of federal agencies within the Department of Health and Human Services and the Department of Agriculture includes support for research on a wide range of public health issues, including obesity prevention. In the private sector, a number of large funders and professional organizations increasingly focus on strategies and policies to prevent obesity in specific demographically defined groups, including children and low-income families. Partnering with public and private organizations and individuals that have supported other multifaceted public health initiatives, such as smoking cessation programs, may also guide or support the crafting of effective communication and dissemination strategies for the framework.

Messages and Materials

The content of messages about the framework—and therefore the materials used to communicate those messages—should take into account the objectives of the communication and dissemination plan, the target audiences, the possible channels for reaching and engaging those audiences, and the resources and goals of potential partners in the communication and dissemination process. The messages delivered should reflect an understanding of the framework's purpose, sensitivity to the current decision-making environment, and understanding of the importance of integrating the framework into current practice. Some materials may be detailed enough to serve as a reference (e.g., this report). Others may briefly underscore the major messages of this report (e.g., its summary or a general report brief). Other materials can target a specific audience; these include targeted report briefs, presentations, training materials, and webinars. These different types of materials can each be customized for a specific audience (e.g., researchers, decision makers, or a general audience) with a specific message (e.g., increase the knowledge base) through a specific channel (e.g., conference, presentation).

Implementation and Monitoring

A detailed implementation plan is essential for effective communication and dissemination of the framework. For example, it is important to indicate how and when resources will be needed, when specific events will occur, and how the communication and dissemination efforts will be assessed. It is also important to determine whether messages are being delivered appropriately, effectively, and efficiently; whether materials are being distributed to the right people and in the right quantities; and whether the intended activities are occurring. Such "process evaluation" helps identify needed corrections to the messages or materials and determine whether the plan is meeting its objectives.

Specific measures can be used to monitor the implementation of the communication and dissemination plan. Such measures could include the following questions: Is the framework or discussion of the framework appearing in peer-reviewed journal articles, grant applications, meeting or conference proceedings, or websites? If so, which ones, and what is their audience? What organizations or associations are using the framework, how are they using it, and who is in their audience network? In what current decision-making activities is the framework being used? How is it being used in the public policy arena? Answers to these and similar questions can be used to modify the messages and materials employed and identify the settings, channels, networks, and members of the target audience that have not been reached.

Finally, to implement and monitor the communication and dissemination plan, it may be necessary to develop a depository, or virtual storehouse, to help track relevant events. It may also be helpful to convene an advisory committee at regular intervals to review the progress being made and communication efforts and their audiences.

EVALUATING AND REFINING THE FRAMEWORK

Changes that occur or are suggested as a result of implementing the framework need to be measured and evaluated to determine how well the framework is meeting its objectives in supporting evidence-generation and decision-making processes and how it could be improved, refined, or integrated to make it more effective.

Objectives

The evaluation and subsequent refinement of the framework have three major objectives. First, it is important to assess the *acceptance* of the framework in the current decision-making environment by measuring its adoption and utilization by the target audiences. Assessing the acceptance of the framework will help determine the degree to which the communication and dissemination plan has achieved its objectives. Second, it is important to measure the *impact* of integrating the framework into the current environment by examining its effect on and thus its relevance to the generation of research and public health policies and programs. Finally, it is essential to *sustain* the use of the framework in current and future evidence-generation and decision-making activities. The results of these evaluations should be used to refine and adapt the framework to incorporate real-world lessons and feedback and sustain its usefulness.

To meet the above objectives, the development of a strategic evaluation and refinement plan should start with defining key questions to address (such as those listed in the next section) and the specific information needed to answer them. These questions could include the outcomes of integrating the framework into evidence-generation and decision-making processes, as well as those of implementing the communication and dissemination plan; exploring the methods and opportunities best suited to capturing and measuring these outcomes; and determining how and when to evaluate these findings and integrate them into the framework.

Desired Outcomes

Key outcomes to be evaluated include the framework's utilization, adoption, acceptance, and impact. The following specific questions should be considered:

- Utilization and adoption
  - Is the framework being incorporated into research?
  - Is it being used by the target audiences or incorporated into their decision-making activities?
  - Is it being used in obesity- and other public health–related prevention research?
- Acceptance
  - Is the framework usable (effective, efficient, and satisfying) for the target audiences (International Organization for Standardization, 1998)?
  - What aspects of the framework are unclear or require additional expertise to understand?
  - Are major concepts of the framework inhibiting its acceptance?
  - Is it feasible for the framework to be used to meet the needs of the target audiences?
- Impact
  - Is the framework changing the way people identify, collect, and generate evidence?
  - Is it changing the way people make decisions?

Data Collection Methods and Opportunities

After specifying the desired outcomes, the next step is to select data collection methods that can best be used to measure them. These methods will most likely include direct observation, questionnaires, or personal interviews. Specific tools may include survey questionnaires, interview guides, training kits, structured interview templates, and tally sheets. The methods and tools used should be based on access to the intended audiences and resources at hand.

A number of opportunities for data collection can be considered. For example, funders of research could request research proposals to study the effects of using the framework. This would have the benefit of encouraging researchers to critique and publish on the progress of the framework's acceptance, use, and integration into public health–related evidence generation and policy making. Activities employed to train the target audiences in the use and integration of the framework could be used to gather immediate feedback, and serve as a source of information over a longer period of time as well. Another opportunity for collecting data on the desired outcomes is monitoring the literature for articles citing the framework.

Integration of Evaluation Findings into the Framework

Evaluating the impact of the framework on public health–related evidence generation and policy making will highlight the framework's strengths and weaknesses. These lessons learned should then be incorporated into the framework to increase its value and sustain its impact.

REFERENCES

HEALCP (Healthy Eating Active Living Convergence Partnership). 2008. *Home*. http://www.convergencepartnership.org (accessed August 21, 2009).

International Organization for Standardization. 1998. *Ergonomic requirements for office work with visual display terminals (VDTS) part II: Guidance on usability*. ISO9241-11. Geneva, Switzerland: International Organization for Standardization.

NCCOR (National Collaborative on Childhood Obesity Research). 2009. *Mission*. http://www.nccor.org/about_mission.html (accessed August 27, 2009).

NCI (National Cancer Institute). 2008. *Making health communication programs work*. Bethesda, MD: U.S. Department of Health and Human Services.

10

Conclusions and Recommendations

Decisions about prevention are complex, not only for the obesity problem but also for other problems with multiple types and layers of causation. Recognition of the need to emphasize population-based approaches to obesity prevention, the urgency of taking action, and the desire of many decision makers to have evidence on which actions to take have created a demand for evidence with which to answer a range of questions. In reality, the evidence approaches that apply to decision making about the treatment of obesity or other clinical problems are inadequate and sometimes inappropriate for application to decisions about public health initiatives. The need to work around evidence gaps and the limitations of using evidence hierarchies that apply to medical treatment for assessing population-based preventive interventions have been faced by the developers of several prior Institute of Medicine (IOM) reports on obesity prevention (focused on child and adolescent obesity). These evidence issues are not new and have already been the focus of many efforts in the field of public health in relation to other complex health problems. However, they are far from resolved. Considering these issues in relation to obesity prevention has the potential to advance the field of public health generally while also meeting the immediate need for clarity on evidence issues related to addressing the obesity epidemic.

The IOM's Food and Nutrition Board formed the Committee on an Evidence Framework for Obesity Prevention Decision Making, with funding from Kaiser Permanente, the Robert Wood Johnson Foundation, and the Centers for Disease Control and Prevention. This committee was asked to develop a framework for evidence-informed decision making in obesity prevention, focused on approaches for assessing policy, environmental, and community interventions designed to influence diet and physical activity. The committee was tasked to:

- provide an overview of the nature of the evidence base for obesity prevention as it is currently construed;

- identify the challenges associated with integrating scientific evidence with broader influences on policy and programmatic considerations;
- provide a practical and action-oriented framework of recommendations for how to select, implement, and evaluate obesity prevention efforts;
- identify ways in which existing or new tools and methods can be used to build a useful and timely evidence base appropriate to the challenges presented by the epidemic, and describe ongoing attempts to meet these challenges;
- develop a plan for communicating and disseminating the proposed framework and its recommendations; and
- specify a plan for evaluating and refining the proposed framework in current decision-making processes.

CONCLUSIONS

Recognition is increasing that overweight and obesity are not only problems of individuals, but also societywide problems of populations. Acting on this recognition will require multifaceted, population-based changes in the socioenvironmental variables that influence energy intake and expenditure. There exist both a pressing need to act on the problem of obesity and a large gap between the type and amount of evidence needed to act and the type and amount of evidence available to meet that need. A new framework is necessary to assist researchers and a broad community of decision makers in generating, identifying, and evaluating the best evidence available and in summarizing it for use in decision making. This new framework also is important for researchers attempting to fill important evidence gaps through studies based on questions with program and policy relevance. However, the methods used and the evidence generated by traditional research designs do not yield all the types of evidence useful to inform actions aimed at addressing obesity prevention and other complex public health challenges. An expanded approach is needed that emphasizes the decision-making process and contextual considerations.

The Framework

To meet this need, the committee developed the L.E.A.D. (**L**ocate Evidence, **E**valuate Evidence, **A**ssemble Evidence, and Inform **D**ecisions) framework, designed to facilitate a systematic approach to the identification, implementation, and evaluation of promising, reasonable actions to address obesity prevention and other complex public health challenges (see Figure 10-1). The framework is designed to help identify the nature of the evidence that is needed and clarify what changes in current approaches to generating and evaluating evidence will facilitate meeting those needs. This section describes the main components of the framework and issues related to these components.

Obesity prevention has not been addressed successfully by traditional study designs, which are generally linear and static. A systems approach is needed to develop more complex, interdisciplinary strategies. Accordingly, the L.E.A.D. framework

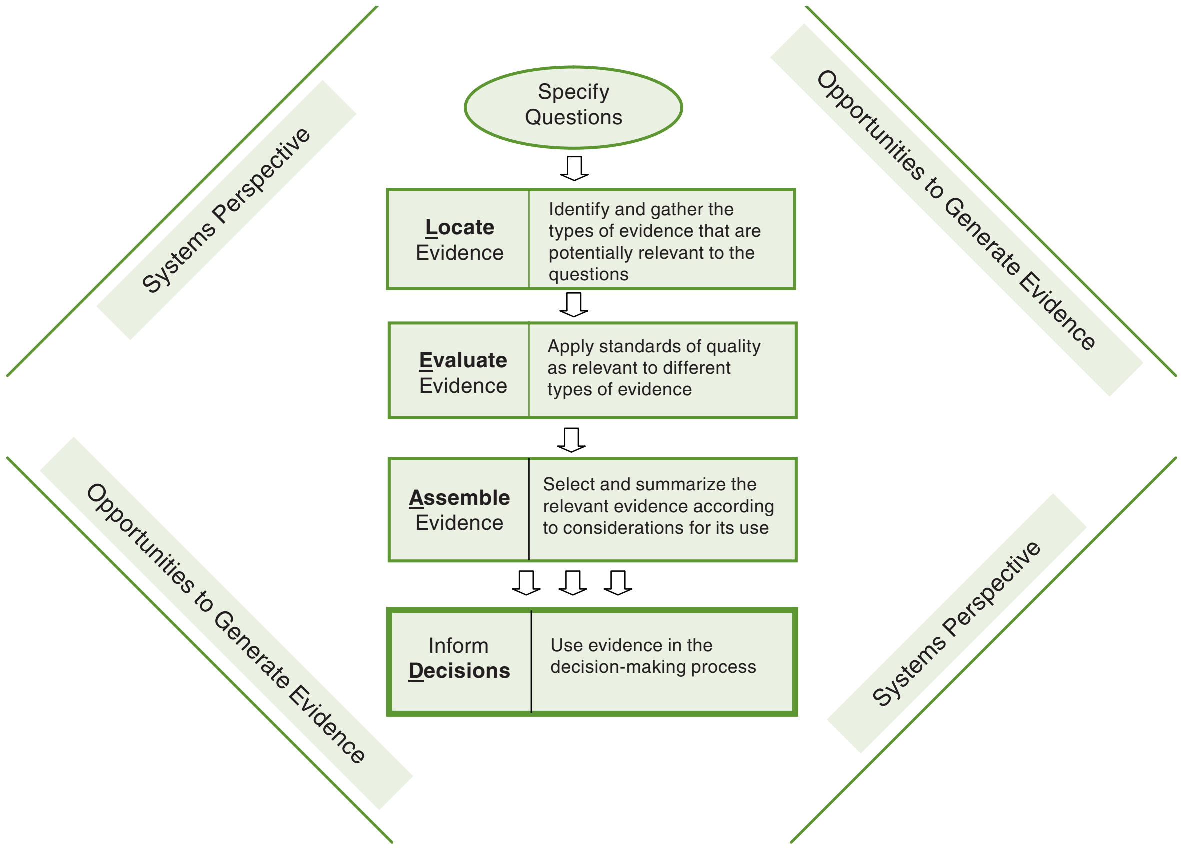


FIGURE 10-1 The **L**ocate Evidence, **E**valuate Evidence, **A**ssemble Evidence, Inform **D**ecisions (L.E.A.D.) framework for obesity prevention decision making.

recommends taking a systems perspective. In other words, it is necessary to use an approach that encompasses the whole picture, highlighting the broader context and interactions among levels, to capture the complexity of obesity prevention and other multifactorial public health challenges.

Addressing such challenges first requires specifying the question(s) being asked to guide the identification of evidence that is appropriate, inclusive, and relevant. Core to the framework is the orientation of the user. A variety of decisions have to be made to address obesity prevention. To capture the resulting mix of evidence needs, the framework adopts a typology that differentiates three broad categories of interrelated questions of potential interest to the user: Why should we do something about this problem? What specifically should we do? and How do we implement this information for our situation? This "Why," "What," "How" typology stresses the need for multiple types of evidence to support decisions on obesity prevention.

Once the question(s) of interest have been specified, locating useful evidence requires clear knowledge of the types of information that may be useful and an awareness of where that information can be found. The framework calls for the use of diverse approaches to gather and synthesize information from other disciplines that address issues similar to those faced in obesity prevention and public health generally. Evidence identified and gathered to inform decision making for obesity prevention and other complex public health challenges should be assessed based on both its generalizability and level of certainty (i.e., its external and internal validity, respectively). The L.E.A.D. framework addresses these two key aspects of the evidence through the nature of the question(s) being asked, established criteria for the value of evidence, and the context in which the question(s) arise. Results of the overall evaluation of evidence should provide answers on what to do, how to do it, and how strongly the action is justified.

When decision makers are coming to a decision on obesity prevention actions, it is important for them to understand the state of the available knowledge relevant to that decision. This knowledge includes evidence on the specific problem to be addressed, the likely effectiveness and impact of proposed actions, and key considerations involved in their implementation. Successful evidence gathering, evaluation, and synthesis for use in obesity prevention usually require the involvement of a number of disciplines using a variety of methodologies and technical languages. The framework incorporates a standardized approach using a uniform language and structure for summarizing the relevant evidence in a systematic, transparent, and transdisciplinary way that is critical for communicating the process and conclusions clearly.

With an emergent problem such as obesity, decisions to act often must be made in the face of a relative absence of evidence, or evidence that is inconclusive, inconsistent, or incomplete. Evidence gathered from a particular intervention implemented in a closely controlled manner within a specific population with its own unique characteristics is often difficult to apply to a similar intervention with another population. The typical way of presenting results of obesity prevention efforts in journals often adds to the problem of incomplete evidence because useful aspects of the research related to its generalizability are not reported. If obesity prevention actions must be taken when evidence is limited, this incomplete evidence can be blended with theory, expert opinion, experience, and local wisdom to make the best decision possible. The actions taken then should undergo critical evaluation, the results of which should be used to build credible evidence for use in decision making about future efforts. Important alternatives to waiting for the funding, implementation, and publication of formal research on obesity prevention are natural experiments as sources of practice-based evidence, "evaluability assessment" of emerging innovations (defined as assessing whether a program is ready for full-scale evaluation), and continuous quality assessment of ongoing programs. The L.E.A.D. framework process leads to knowledge integration, or the incorporation of new knowledge gained through the process of applying the framework into the context of the organization or system where decisions are made.

The evidence base to support the identification of effective obesity prevention interventions is limited in many areas. Opportunities to generate evidence may occur at any phase of the evidence review or decision-making process. The L.E.A.D. framework guides the generation of evidence related to "What," "Why," and "How" questions and supports the use of multiple forms of evidence and research designs from a variety of disciplines. In obesity prevention–related research, the generation of evidence from evaluation of ongoing and emerging initiatives is a particular priority.

Next Steps

Researchers, decision makers, and intermediaries working on obesity prevention and other complex multifactorial public health problems are the primary audiences for communicating and disseminating the L.E.A.D. framework. With sufficient information, they can apply the framework as a guide for generating needed evidence and supporting decision making. It is important to understand the settings, communication channels, and activities of these key audiences to engage and educate them effectively on the purpose and adoption of the framework. To support the development of a communication and dissemination plan, it is critical to create partnerships, make use of existing activities and networks, and tailor the messages and approaches to each target audience.

As the target audiences begin to use the framework, assessing its use in selected settings will be essential so it can be improved and refined. Evaluation of the impact of the L.E.A.D. framework is also important for determining its relevance to current evidence-generation and decision-making processes. To this end, key outcome measures—utilization, adoption, acceptance, maintenance, and impact—should be defined and data collected on these measures. It will be important to develop or adopt data collection tools and utilize methods and existing initiatives that will best serve this purpose, as well as to systematically integrate the feedback thus obtained to sustain and improve the framework's applicability and utilization.

RECOMMENDATIONS

The United States has made progress toward translating science into practice in the brief time since the obesity epidemic was officially recognized. But the pace of this translation has been slow relative to the scope and urgency of the problem and the associated harms and costs. As discussed above, moreover, the evidence emerging from applied research on obesity prevention can be inconclusive, incomplete, and inconsistent. A systematic process is needed to improve the use of available evidence and increase and enhance the evidence base to inform decisions on obesity prevention and other complex public health problems. Commitment to such a process is needed from both decision makers and those involved in generating evidence, including public and private policy makers and their advisors, scientific and policy think tanks, advocacy groups and stakeholders, program planners, practitioners in public health and other sectors, program evaluators, public health researchers and research scientists, journal editors, and funders. With this in mind, the committee makes the following recommendations for assisting decision makers and researchers in using the current evidence base for obesity prevention and for taking a systems-oriented, transdisciplinary approach to generate more, and more useful, evidence.

Utilize the L.E.A.D. Framework

Recommendation 1: Decision makers and those involved in generating evidence, including researchers, research funders, and publishers of research, should apply the L.E.A.D. framework as a guide in their utilization and generation of evidence to support decision making for complex, multifactorial public health challenges, including obesity prevention.

Key assumptions that should guide the use of the framework include the following:

- A systems perspective can help in framing and explaining complex issues.
- The types of evidence that should be gathered to inform decision making are based on the nature of the questions being asked, including Why? ("Why should we do something about this problem in our situation?"), What? ("What specifically should we do about this problem?"), and How? ("How do we implement this information for our situation?"). A focus on subsets of these questions as a starting point in gathering evidence explicitly expands the evidence base that is typically identified and gathered.
- The quality of the evidence should be judged according to established criteria for that type of evidence.
- Both the level of certainty of the causal relationship between an intervention and the observed outcomes and the intervention's generalizability to other individuals, settings, contexts, and time frames should be given explicit attention.
- The analysis of the evidence to be used in making a decision should be summarized and communicated in a systematic, transparent, and transdisciplinary manner that uses uniform language and structure. The report on this analysis should include a summary of the question(s) asked by the decision maker; the strategy for gathering and selecting the evidence; an evidence table showing the sources, types, and quality of the evidence and the outcomes reported; and a concise summary of the synthesis of selected evidence on why an action should be taken, what that action should be, and how it should be taken.
- If action must be taken when evidence is limited, this incomplete evidence can be blended carefully and transparently with theory, expert opinion, and collaboration based on professional experience and local wisdom to support making the best decision.

Sustained commitments will be needed from both the public and private sectors to achieve successful utilization of the various elements of the L.E.A.D. framework in future evidence-informed decision making and evidence generation. This responsibility lies with the academic and research community, as well as with government and private funders and the leadership of journals that publish research in this area. Necessary supports will include increasing understanding of systems thinking and incorporating it into research-related activities, creating and maintaining resources to support the utilization of evidence, establishing standards of quality for different types of evidence, and supporting the generation of evidence, each of which is described in more detail below. Finally, it will be necessary to communicate, disseminate, evaluate, and refine the L.E.A.D. framework.

Incorporate Systems Thinking

Recommendation 2: Researchers, government and private funders, educators, and journal editors should incorporate systems thinking into their research-related activities.

To implement this recommendation:

- Researchers should use systems thinking to guide the development of environmental and policy interventions and study designs.
- Government and private funders should encourage the use of systems thinking in their requests for proposals and include systems considerations in proposal evaluations.
- Universities, government agencies such as the U.S. Centers for Disease Control and Prevention, and public health organizations responsible for educating public health practitioners and related researchers should establish training capacity for the science and understanding of systems thinking and the use of systems mapping and other quantitative or qualitative systems analysis tools.
- Journal editors should encourage the use of systems thinking for addressing complex problems by developing panels of peer reviewers with expertise in this area and charging them with making recommendations for how authors could use systems thinking more effectively in their manuscripts.

Build a Resource Base

Recommendation 3: Government, foundations, professional organizations, and research institutions should build a system of resources (people, compendiums of knowledge, registries of implementation experience) to support evidence-based public policy decision making and research for complex health challenges, including obesity prevention.

To implement this recommendation:

- The Secretary of Health and Human Services, in collaboration with other public- and private-sector partners, should establish a sustainable registry of reports on evidence for environmental and policy actions for obesity prevention.

Integral to this registry should be the expanded view of evidence for decision making on obesity prevention proposed in this report and the sharing of experiences and innovative programs as the evidence evolves. A service provided by this registry should be periodic synthesis reviews based on mixed qualitative and quantitative methods.

- The Secretary of Health and Human Services, in collaboration with other public- and private-sector partners, should develop and fund a resource for compiling and linking existing databases that may contain useful evidence for obesity prevention and related public health initiatives. This resource should include links to data and research from disciplines and sectors outside of obesity prevention and public health and to data from nonacademic sources that are of interest to decision makers.

Establish Standards for Evidence Quality

Recommendation 4: Government, foundations, professional organizations, and research institutions should catalyze and support the establishment of guidance on standards for evaluating the quality of evidence for which such standards are lacking.

To implement this recommendation:

- Government and private funders should give priority to funding for the development of guidance on standards for evaluating the quality of the full range of evidence types discussed in this report that are useful in making obesity prevention decisions, especially those for which the scientific literature is limited.
- Professional organizations and research institutions should encourage and bring attention to efforts by faculty, researchers, and students to establish guidance in this area.

Support the Generation of Evidence

Recommendation 5: Obesity prevention research funders, researchers, and publishers should consider, wherever appropriate, the inclusion in research studies of a focus on the generalizability of the findings and related implementation issues at every stage, from conception through publication.

To implement this recommendation:

- Those funding research in obesity prevention should give priority to support for studies that include an assessment of the limitations, potential utility, and applicability of the research beyond the particular population, setting, and circumstances in which the studies are conducted, including by initiating requests for applications and similar calls for proposals aimed at such studies. Additional ways in which this recommendation could be implemented include adding criteria related to generalizability to proposal review procedures and training reviewers to evaluate generalizability.

- Obesity prevention researchers and program evaluators should give special consideration to study designs that maximize evidence on generalizability.
- Journal editors should provide guidelines and space for authors to give richer descriptions of interventions and the conditions under which they are tested to clarify their generalizability.

Recommendation 6: Research funders should increase opportunities for those carrying out obesity prevention initiatives to measure and share their outcomes so others can learn from their experience.

To implement this recommendation:

- Organizations funding or sponsoring obesity prevention initiatives—including national, regional, statewide, or local programs; policy changes; and environmental initiatives—should provide resources for obtaining practice-based evidence from innovative and ongoing programs and policies in a more routine, timely, and systematic manner to capture their processes, implementation, and outcomes. These funders should also encourage and support assessments of the potential for evaluating the most innovative programs in their jurisdictions and sponsor scientific evaluations where the opportunities to advance generalizable evidence are greatest.
- Research funders, researchers, and journal editors should assign higher priority to studies that test obesity prevention interventions in real-world settings in which major contextual variables are identified and their influence is evaluated.

Recommendation 7: Research funders should encourage collaboration among researchers in a variety of disciplines so as to utilize a full range of research designs that may be feasible and appropriate for evaluating obesity prevention and related public health initiatives.

To implement this recommendation:

- As part of their requests for proposals on obesity prevention research, funders should give priority to and reward transdisciplinary collaborations that include the creative use of research designs that have not been extensively used in prevention research but hold promise for expanding the evidence base on potential environmental and policy solutions.

Communicate, Disseminate, Evaluate, and Refine the L.E.A.D. Framework

Recommendation 8: A public–private consortium should bring together researchers, research funders, publishers of research, decision makers, and other stakeholders to discuss the practical uses of the

L.E.A.D. framework, and develop plans and a timeline for focused experimentation with the framework and for its evaluation and potential refinement.

To implement this recommendation:

- Interested funders should bring together a consortium of representatives of key stakeholders (including decision makers, government funders, private funders, academic institutions, professional organizations, researchers, and journal editors) who are committed to optimizing the use of the current obesity prevention evidence base and developing a broader and deeper base of evidence.
- This consortium should develop an action-oriented plan for funding and implementing broad communication, focused experimentation, evaluation, and refinement of the L.E.A.D. framework. This plan should be based on the major purposes of the framework: to significantly improve the evidence base for obesity prevention decision making on policy and environmental solutions, and to assist decision makers in using the evidence base.

A

Acronyms and Glossary

ACRONYMS

AIDS	Acquired immunodeficiency syndrome
ASSIST	American Stop Smoking Intervention Study for Cancer Prevention
BMI	Body mass index
CBO	Congressional Budget Office
CBPR	Community-based participatory research
CDC	U.S. Centers for Disease Control and Prevention
COCOMO	Common Community Measures for Obesity Prevention
COMMIT	Community Intervention Trial for Smoking Cessation
CONSORT	Consolidated Standards of Reporting Trials
EBM	Evidence-based medicine
EBPH	Evidence-based public health
EBPP	Evidence-based public policy
FCC	Federal Communications Commission
GHG	Greenhouse gas
GRADE	Grading of Recommendations Assessment, Development, and Evaluation
HEALCP	Healthy Eating Active Living Convergence Partnership
HHS	U.S. Department of Health and Human Services
HIV	Human immunodeficiency virus

IOM	Institute of Medicine
IOTF	International Obesity Task Force
ISIS	Initiative on the Study and Implementation of Systems
ITS	Interrupted time series
L.E.A.D.	Locate Evidence, Evaluate Evidence, Assemble Evidence, Inform Decisions
MATCH	Multilevel Approach to Community Health
MIT	Massachusetts Institute of Technology
NCCOR	National Collaborative on Childhood Obesity Research
NCI	National Cancer Institute
NHANES	National Health and Nutrition Examination Survey
NIH	National Institutes of Health
PATCH	Planned Approach to Community Health
RCT	Randomized controlled trial
RD	Regression discontinuity
REACH	Racial and Ethnic Approaches to Community Health
RE-AIM	Reach, Effectiveness, Adoption, Implementation, Maintenance
RWJF	The Robert Wood Johnson Foundation
SAMHSA	Substance Abuse and Mental Health Services Administration
SD	Standard deviation
SERCA	Special Emphasis Research Career Award
SSB	Sugar-sweetened beverage
SUTVA	Stable unit treatment value assumption
TREND	Transparent Reporting of Evaluations with Nonrandomized Designs
USDA	U.S. Department of Agriculture
USPSTF	United States Preventive Services Task Force
WHO	World Health Organization
WIC	Special Supplemental Nutrition Program for Women, Infants, and Children

GLOSSARY

Adaptability

The ability to modify an intervention to fit the cultural context in which the intervention will take place and the unique circumstances of the agency and other stakeholders, without contradicting the intervention's core elements and internal logic.

Adoption

The decision to use an intervention as the best course of action available. Within the RE-AIM framework, adoption is measured as the absolute number, proportion, and representativeness of settings and/or intervention staff (people who deliver the program) that are willing to initiate a program.

Archival study

The use of a model based on past evidence or decisions on a behavior or intervention for purposes of predicting future behavior.

Assessment

In this report, refers to the process of observing, describing, collecting data on, and measuring the quality and effectiveness of an intervention or policy. See also *Evaluation*.

Benefits

Positive effects brought about by an intervention or policy (intended or unintended).

Body mass index

An indirect measure of body fat, calculated as the ratio of a person's body weight (in kilograms) to the square of a person's height (in meters):

$$\text{BMI (kg/m}^2) = \text{weight (kilograms)} \div \text{height (meters)}^2$$

$$\text{BMI (lb/in}^2) = \text{weight (pounds)} \div \text{height (inches)}^2 \times 703$$

In adults, a body mass index (BMI) of 25 or greater is considered overweight, and a BMI of 30 or greater is considered obese. In children and youth, BMI is based on growth charts for age and gender and is referred to as BMI for age, which is used to assess underweight, overweight, and risk for overweight. According to the Centers for Disease Control and Prevention (CDC), a child with a BMI for age that is equal to or greater than the 95th percentile is considered obese. A child with a BMI for age that is equal to or between the 85th and 95th percentiles is overweight.

BMI *z*-score

Number of standard deviations away from the population mean BMI; in other words, the degree to which an individual's measurement deviates from what is expected for that individual.

Case study research

Qualitative research in which the investigator explores a bounded system (a case) or multiple bounded systems (cases) over time through detailed, in-depth data collection involving multiple sources of information (e.g., observations, interviews, audiovisual material, and documents and reports), and reports a case description and case-based themes.

Case-control or case-comparison study, retrospective study

Quantitative research in which persons with a disease (or other outcome) of interest and a suitable control group of persons without the disease are studied. The relationship of an attribute to the disease is examined by comparing the diseased and nondiseased with regard to how frequently the attribute is present. Risk is estimated by the odds ratio.

Causality

The relationship of causes to the effects they produce. A cause is termed "necessary" when it must always precede an effect; this effect need not be the sole result of the one cause. A cause is termed "sufficient" when it inevitably initiates or produces an effect. Any given causal factor may be necessary, sufficient, neither, or both.

Cluster randomization design

An experiment in which clusters of individuals rather than independent individuals are randomly allocated to the intervention group.

Cohort study

A method of study in which subsets of a defined population can be identified who are, have been, or in the future may be exposed or not exposed, or exposed in different degrees, to a factor or factors hypothesized to influence the probability of occurrence of a given disease or other outcome. The main feature of cohort studies is observation of large numbers of individuals over a long period (commonly years), with comparison of incidence rates in groups that differ in exposure levels. Risk is estimated by the relative risk.

Community

A social entity that can be either spatial (based on where people live in local neighborhoods, residential districts, or municipalities) or relational (based on common ethnic, cultural, or other characteristics or similar interests).

Community assessment (needs assessment)

A procedure used to collect data that describe the needs and strengths of a specific group, community, or population.

Comparative effectiveness research

Research involving the generation and synthesis of evidence that compares the benefits and harms of alternative methods for preventing, diagnosing, treating, and monitoring health conditions in a real-world setting.

Comparative experiment

A study that compares the effects of a program or policy against a comparison condition or control group.

Content analysis

A systematic and objective summary of the concepts of text.

Context

The set of factors or circumstances that surrounds a situation or event and gives meaning to its interpretation; the broader environment in which a program operates.

Context evaluation

See *Needs assessment*.

Cost/benefit analysis

Used to compare the various costs associated with an investment in a program or initiative with its proposed benefits.

Cost-effectiveness analysis

Used to assess the most efficient method for achieving a program or policy goal. The costs of alternatives are measured in terms of their requisite estimated dollar expenditures. Effectiveness is defined by the degree of achievement of a goal and may be measured in dollars.

Cost-feasibility analysis

A method used to determine whether an option or intervention is affordable.

Cost-minimization analysis

A method used to assess the least costly way of achieving the same outcome in which consequences of competing interventions are the same, and only costs are taken into consideration.

Cost-utility analysis

Used to assess the relative economic value (e.g., cost-utility ratio) of alternative strategies aimed at achieving similar outcomes. This type of analysis converts effects into personal preferences and describes how much it costs for some additional quality gain (e.g., cost per additional quality-adjusted life-year).

Cross-sectional study or survey, prevalence study

Quantitative research involving the observation of a defined population at a single point in time or time interval.

Decision maker

In this report, refers to any individual who makes a decision to implement an intervention affecting obesity prevention or another public health problem. Examples of decision makers are state legislators, mayors, public health officials, city council members, school board members, employers, and community organization officials.

Disparities

A term used to describe differences in quality of health and health care across racial, ethnic, and socioeconomic groups.

Dissemination

In the context of this report, the targeted distribution of information and intervention materials to a specified audience. The intent is to diffuse knowledge and the associated evidence-based interventions.

Ecological model

A model that suggests the importance of addressing individual, interpersonal, organizational, community (including social and economic features), and health policy factors because of their effect on individual behavior change and their direct effect on health.

Economic cost analysis

A study that assesses the direct and indirect costs associated with a health problem or intervention.

Effect

The change in an outcome that results from an intervention.

Effect size

The estimated magnitude of an effect.

Effectiveness

The extent to which an intended effect or benefits that could be achieved under optimal conditions are also achieved in practice.

Efficacy

The extent to which an intervention can be shown to be beneficial under optimal conditions.

Energy balance

A state in which caloric intake is equivalent to energy expenditure, resulting in no net weight gain or weight loss.

Energy expenditure

Calories used to support the body's basal metabolic needs plus those used for themogenesis, growth, and physical activity.

Energy intake

Calories ingested as foods and beverages.

Evaluability assessment

A method of determining whether activities necessary for a successful full-scale evaluation are in place, whether the activities include program objectives that are well defined and plausible, whether a program can realistically achieve its intended goals, and whether the intended uses of information from evaluation have been clearly specified.

Evaluation

Systematic assessment of the quality and effectiveness of an intervention or policy and its effects to produce information that can be used by those with an interest in the intervention's improvement or effectiveness.

Evidence-based medicine

The conscientious, explicit, and judicious use of current best evidence in making decisions about the care of individual patients. The practice of evidence-based medicine means integrating individual clinical expertise with the best available external clinical evidence from systematic research.

Evidence-based public health

The development, implementation, and evaluation of effective programs and policies in public health through the application of principles of scientific reasoning, including systematic uses of data and information systems and appropriate use of behavioral science theory and program planning models.

Evidence farming

A flexible, infrastructural digital tool that enables systematic attention to, acquisition of, and utilization of local experience to improve clinical practice and patient outcomes.

Evidence synthesis

An approach that involves the application of techniques for combining multiple sources of evidence.

Experimental design

A study that explores the effects of manipulating, allocating, and timing an intervention.

Expert knowledge (opinion)

The views of professionals with expertise in a particular form of practice or field of inquiry, such as clinical practice or research methodology. Expert knowledge may refer to one person's views or to the consensus view of a group of experts.

External validity

The property that permits research results to be generalized to other population units, including other individuals, settings, contexts, and time frames, and thereby to inform practice.

Feasibility

The likelihood of successful implementation of an intervention.

Fidelity

The degree to which an intervention was implemented according to protocol, as intended, or as planned. Also referred to as "integrity" of implementation.

Focus group

A qualitative research method whereby a moderator convenes a group of participants who often have common characteristics (e.g., age, gender, ethnicity) to discuss the attributes of a particular concept or product. Focus groups are often used in the marketing development phase to generate ideas and provide insight into consumer reactions and perceptions.

Framework

Defined in this report to mean a process for identifying, evaluating, and compiling evidence relevant to various potential decision-making scenarios for obesity prevention and other complex public health problems based on clearly specified concepts, assumptions, or rules.

Frequency

The number of observations of a variable under investigation. Frequency of a disease or risk factor as discussed in this report refers to obesity or obesity-related diseases.

Generalizability

The extent to which the results produced by a particular intervention or set of interventions under specific conditions can be expected to produce similar findings in different settings or contexts with different populations. See also *External validity*.

Geographic mapping

A method used to identify graphically the location of an outcome or variable of interest.

Gray literature

Research reports that are not found in traditional peer-reviewed publications but in, for example, government agency monographs, symposium proceedings, and unpublished company reports.

Guidelines

In the present context, standardized information describing the best practices for addressing health problems commonly encountered in public health practice. The information is based on scientific evidence for the effectiveness and efficiency of the practices described. Where such evidence is lacking, guidelines are sometimes based on the consensus opinions of public health experts.

Harm

Adverse effects brought about by an intervention.

Health disparities

Differences in the occurrence, frequency, mortality, and burden of diseases and other unfavorable health conditions that exist among specific population groups, including racial and ethnic minorities.

Health impact assessment

A combination of procedures, methods, and tools by which a policy, program, or project can be evaluated in terms of its potential effects on the health of a population and the distribution of those effects within the population.

Impact

The potential effects of an intervention on the health of a population.

Impact evaluation (or assessment)

A measure of whether an outcome is the result of an intervention. To support a cause-and-effect conclusion, an impact evaluation incorporates research methods that eliminate alternative explanations for an outcome and shows how much difference an intervention can make compared with the effect seen when no intervention is used. In public health programs, an impact evaluation often occurs over a long period of time.

Implementation

The act of putting an intervention into practice. Within the RE-AIM framework, implementation is measured in terms of the intervention agents' fidelity to the various elements of the intervention's protocol, including consistency of delivery as intended and time and cost.

Incidence

The frequency of new cases of a condition or disease within a defined time period. Incidence is commonly measured in terms of new cases per 1,000 (or 100,000) population at risk per year.

Indicated intervention

A preventive intervention that targets high-risk individuals, in the present context those identified as having a minimal but detectable amount of excess weight that foreshadows obesity.

Internal validity

The property that permits an observed effect to be attributed with certainty to a specific intervention, allowing for inferences about causality or a link between a given intervention and client/patient outcomes. Randomization is viewed as one of the best means of maximizing internal validity.

Interrupted time series study

A quantitative study in which the initial observations serve as a control, and after an intervention is introduced, the remaining observations are experimental.

Intervention

A policy, program, or action intended to bring about identifiable outcomes.

Knowledge integration

The effective incorporation of knowledge into the decisions, practices, and policies of organizations and systems.

Knowledge translation

An encompassing term that denotes the exchange, synthesis, and ethically sound application of research findings within a complex system of relationships among researchers and knowledge users; the incorporation of research knowledge into policies and practices, thus translating knowledge into improved health of the population.

Logic model

A diagram that illustrates the sequencing of program activities that should occur for planning, organizing, implementing, and producing desired results.

Longitudinal study, cohort study, prospective study

A study that examines quantitatively the specific characteristics of individuals, subgroups, or populations over time.

Market research

Research involving the gathering of information about markets and customers. Market research can include information on target markets and their needs, competitors, market trends, and customer satisfaction with products and services.

Matched cohort study

A cohort design in which the groups are matched on many variables (e.g., economic or health status) so the variable of interest (e.g., BMI) can be isolated.

Matching, mapping, pooling, and patching

The use of theory, professional experience, and indigenous wisdom to adapt evidence to local settings, populations, and times.

Mediator

The mechanism by which one variable affects another.

Meta-analysis

A systematic, quantitative method for combining information from multiple studies to derive the most meaningful answer to a specific question.

Mixed-method design

The use of methodologies drawn from a variety of disciplines. The research employs both qualitative and quantitative data gathering and analysis methods that combine extensive descriptions of context and the experiences of program participation with standardized assessments of changes in institutions or systems, the environment, and individual or population behaviors.

Moderator

A variable that changes the impact of one variable on another.

Monitoring

The collection and analysis of data on an intervention or policy as it progresses to ensure the integrity of its planned implementation.

Multifaceted intervention

An intervention involving multiple components that are delivered concurrently to a target group in combination.

Multilevel intervention

An intervention that targets groups at multiple levels of an ecological model.

Multisector intervention

An intervention that mobilizes multiple sectors of society, such as government, civil society, the public, and the private sector.

Narrative review

A process by which a body of literature is collected, reviewed, and summarized using unsystematic methods.

Natural experiments

Naturally occurring circumstances in which different populations are exposed or not exposed to a potentially causal factor such that the circumstances resemble a true experiment in which study participants are assigned to exposed and unexposed groups.

Needs assessment

An evaluative study that answers questions about the social conditions a program is intended to address and the need for the program.

Nonrandomized experiment

See *Observational study, Quasi-experimental study*.

Obesity

See *Body mass index*.

Observational study (nonexperimental)

A design in which variables are observed rather than manipulated.

Outcome

The changes that result from an intervention's activities and outputs. Depending on the nature of an intervention and the theory of change that guides it, an outcome can be short-, intermediate-, or long-term. Indicators or indices are used to assess whether progress has been made toward achieving specific outcomes as a result of an intervention.

Outcome evaluation

An approach to assessing whether anticipated changes have occurred as a result of an intervention. This type of evaluation assesses the extent of change in targeted attitudes, values, behaviors, policies, programs, environments, or conditions between the baseline measurement and subsequent points of measurement over time.

Overweight

See *Body mass index*.

Parallel evidence

Evidence of the effectiveness of an intervention that is being considered to address a problem such as obesity prevention in addressing another public health issue using similar strategies.

Policy

A written statement reflecting a plan or course of action of a government, business, community, or institution that is intended to influence and guide decision making. For a government, a policy may consist of a law, regulation, ordinance, executive order, or resolution.

Policy analysis

An interdisciplinary field that involves research into the nature of policy and its types, the making of policies, and their implications and consequences.

Policy maker

In this report, refers to a decision maker who has the authority to pass legislation affecting obesity prevention or other public health problems. Examples of policy makers are federal or state legislators, governors, and mayors.

Policy scan, policy tracking

A quantitative survey of recommendations and actions, as well as facts, trends, and opinions, related to a specific issue over a predetermined period of time.

Population-based

A term used to describe obesity prevention interventions aimed primarily at the macro level, for example, initiatives that are undertaken to improve environments for eating and physical activity in ways that will reduce the likelihood of excess weight gain among people in general.

Population health

The state of health of an entire community or population as opposed to that of an individual. It is concerned with the interrelated factors that affect the health of populations over the life course and the distribution of patterns of health outcomes.

Practice-based interventions (or evidence)

Interventions, or evidence derived therefrom, that have been developed, implemented, and documented in practice but have not been tested in a more formal, research study.

Predictive studies
The use of models to identify or predict more accurately a behavior or other outcome of interest.

Pretest–posttest design
A quasi-experimental study design that includes a baseline measure of the outcome of interest (pretest), the treatment, and a posttest measure of the outcome to evaluate the effects of an intervention.

Prevalence
The number of instances of a condition or disease in a population at a designated period of time, usually expressed as a percentage of the total population.

Prevention
With regard to obesity, primary prevention represents avoiding the occurrence of obesity in a population, secondary prevention represents the early detection of disease through screening with the purpose of limiting its occurrence, and tertiary prevention involves preventing the sequelae of obesity in childhood and adulthood.

Program
An integrated set of planned strategies and activities that support clearly stated goals and objectives and lead to desirable changes and improvements in the well-being of people, institutions, environments, or all of these.

Qualitative data (or research)
Nonnumerical observations made using approved methods such as participant observation, group interviews, or focus groups. Qualitative data can enrich the understanding of complex problems and help explain why things happen.

Quantitative data (or research)
Data in numerical quantities, such as continuous measurements or counts.

Quasi-experimental study
A quantitative experiment in which the investigator lacks full control over the allocation or timing of an intervention.

Randomized controlled trial
A study design in which groups of patients are randomized to an experimental and a control condition. The groups are monitored for the variables or outcomes of interest.

Randomized encouragement design
An experiment in which treatment assignment is encouraged instead of enforced because of ethical or feasibility issues.

Randomized study
An experiment in which units (i.e., individuals, groups, communities) are assigned randomly to conditions.

Reach
The absolute number, proportion, and representativeness of individuals who participate in an intervention. Within the RE-AIM framework, reach is measured as the absolute number, proportion, and representativeness of individuals who are willing to participate in a given intervention.

RE-AIM framework

A systematic way of reporting research results and translating research into practice. RE-AIM stands for reach, effectiveness, adoption, implementation, and maintenance.

Realist review

A relatively new strategy for synthesizing research that has an explanatory rather than a judgmental focus. The aim is to understand how complex programs work (or why they fail) in particular contexts and settings. Realism is a methodological orientation that has roots in other fields, but is as yet largely untried as an approach to the synthesis of evidence in health care.

Reductionist research

A model that approaches statements of causality by isolating, simplifying, and holding constant key conditions as an attempt is made to understand effects by controlling or removing all potential confounders.

Regression discontinuity design

A study design that requires assignment of participants to treatment or control groups based on whether they fall above or below a cutoff score on a variable measured prior to the treatment.

Representativeness

A component of external validity. With respect to settings, one should ask whether comparisons are made of the similarity of settings under study to the target program settings or to those settings that decline to participate; with respect to individuals, one should ask whether analyses are conducted of the similarities and differences between people who participate in a study and either the target audience or those who decline to participate.

Safety

The condition of being either protected from or unlikely to cause danger, risk, or injury that may be either perceived or objectively defined.

Scientific literature

The body of theoretical and research publications in scientific journals, reference books, textbooks, government reports, policy statements, and other materials on the theory, practice, and results of scientific inquiry.

Secondary analysis

A method used to analyze an existing pool of data to pursue a research interest that differs from that of the original work.

Sector

A distinct subset of a market, society, industry, or the economy in which the members share similar characteristics. Examples of the sectors discussed in this report include government or the public sector, communities, nonprofit and philanthropic organizations, health care, business or the private sector, schools, and households.

Selective intervention

A preventive intervention targeting a subgroup of the population whose risk of developing obesity is above average or high.

Sensitivity analysis

A technique for systematically changing parameters in a model to determine the effects of such changes. This technique is useful when one is attempting to determine the impact on the actual outcome if a particular parameter differs from what was previously assumed.

Simulation modeling

A tool to help one understand the interaction of the parts of a system and a system as a whole. A model (a simplified representation of an actual system) is manipulated in such a way that it operates on time or space to compress it, allowing one to see the interactions that would not otherwise be apparent because of their separation in time or space.

Social determinants

The features and pathways by which societal conditions affect health and can potentially be altered by action. Social determinants discussed in this report include environmental conditions that contribute to disease or risk factors.

Stakeholder

An individual or organization with an interest in an intervention, policy, or outcome.

Strategic planning

In the context of this report, detailed plans that include objectives and a set of essential actions (preventive, therapeutic) believed sufficient to control a health problem.

Survey

An investigation in which information is systematically collected but the experimental method is not used. The information collected almost always requires editing, coding, data entry, and data analysis. Survey data differ from surveillance data in that they are not ongoing but sporadic.

Sustainability

The likelihood that a program will continue over a period of time after grant funding has ended; in broader terms, society's ability to shape its economic and social systems to maintain both natural resources and human life. Sustainability improves quality of life in the present while ensuring continued prosperity in the future.

System dynamics

A methodology for mapping and modeling the forces of change in a complex system to better understand their interaction and govern the direction of the system. It enables stakeholders to combine input to form a dynamic hypothesis that uses computer simulation to compare various scenarios for achieving change.

Systematic review

A review of a clearly formulated question that uses systematic and explicit methods to identify, select, and critically appraise relevant research and to collect and analyze data from the studies that are included in the review. Statistical methods (meta-analysis) may or may not be used to analyze and summarize the results of the included studies.

Systems approach
A paradigm or perspective involving a focus on the whole picture and not just a single element, awareness of the wider context, an appreciation for interactions among different components, and transdisciplinary thinking.

Systems investigation
A promising new frontier for research and action in response to complex and critical challenges, based on a systems approach.

Systems perspective
See *Systems approach*.

Systems science
Research related to systems theory that offers insight into the nature of the whole system that often cannot be gained by studying the system's component parts in isolation.

Systems theory
An interdisciplinary theory that requires a merging of multiple perspectives and sources of information and deals with complex systems in technology, society, and science.

Systems thinking
An iterative learning process in which one takes a broad, holistic, long-term perspective on the world and examines the linkages and interactions among its elements.

Target population
A group of individuals at risk targeted by an intervention.

Transdisciplinary
Researchers working jointly using a shared conceptual framework drawing together disciplinary-specific theories, concepts, and approaches to address a common problem.

Transferability
The generalizability or applicability of inferences obtained in a study to other individuals or entities, settings or situations, time periods, or methods of observation and measurement. See also *External validity*, *Generalizability*, and *Knowledge translation*.

Translation
The process of exchange between research and practice to foster the widespread dissemination of research results and enhance policy and practice initiatives. See also *Knowledge translation*.

Trend
The rate of increase of disease or determinants over time.

Universal intervention
A preventive intervention designed for everyone in the eligible population.

Validity
The extent to which an instrument directly and accurately measures what it is intended to measure.

Variable
Anything that is not constant but that can and does change in different circumstances.

B

Other Evidence Projects

Table B-1 lists a sampling of current efforts to improve research and practice in the field of obesity prevention or public health. The committee considered these efforts when assessing the strengths and limitations of current approaches for generating and evaluating evidence on obesity prevention. The efforts listed in this table are focused on the provision of guidelines for selecting prevention interventions, identifying appropriate outcome measures for obesity prevention, evaluating the strength of evidence for prevention approaches, and using evidence. Sources of systematic reviews are also included.

TABLE B-1 Other Evidence Projects

Project Name	Funder(s)/Investigator(s)	Purpose	Website
Guidelines			
Guide to Community Preventive Services	Centers for Disease Control and Prevention (CDC)	To inform the choice of programs and policies for improving health and preventing disease in a given community.	http://www.thecommunityguide.org
Guidelines, Development, and Recommendations Team	CDC's Division of Nutrition, Physical Activity, and Obesity	To establish a scientific evidence base for what works. This team works in conjunction with the *Guide to Community Preventive Services* and the U.S. Task Force on Community Preventive Services to issue recommendations for obesity prevention and control interventions in worksite, medical care, and community settings, and to provide materials that will help users implement the interventions.	None

continued

TABLE B-1 Continued

Project Name	Funder(s)/Investigator(s)	Purpose	Website
Measures			
Common Community Measures for Obesity Prevention (COCOMO)	CDC, CDC Foundation, Robert Wood Johnson Foundation (RWJF), W.K. Kellogg Foundation, Kaiser Permanente	To identify a core set of community-level measures, focusing on policy and environmental strategies that impact obesity prevention that can be used for research and can serve as baseline variables for planning and evaluation.	http://www.cdc.gov/mmwr/preview/mmwrhtml/rr5807a1.htm
Evaluation			
Center of Excellence for Training and Research Translation	University of North Carolina, CDC	To identify, translate, and disseminate evidence-based interventions, best practices/processes, and implementation tools for use by public health practitioners.	http://www.center-trt.org/
Early Assessment of Programs and Policies to Prevent Childhood Obesity	CDC Division of Adolescent and School Health, RWJF, Prevention Research Centers Program Office, CDC Foundation, and Macro International Inc.	To identify a set of promising local programs and policies (addressing the physical, social, or economic environment) implemented to address obesity among children aged 3-17 and to determine which ones merit rigorous evaluation. Priority is placed on programs and policies implemented in community settings and targeting low-income children to improve eating habits and physical activity levels.	http://www.cdc.gov/prc/research-projects/associated-projects/identifying-local-childhood-obesity-prevention-programs-policies.htm
Practice-Based Evidence Project in Worksites—Swift Worksite Assessment and Translation (SWAT)	CDC Division of Nutrition, Physical Activity, and Obesity	To judge the promise of worksite health promotion practices. The SWAT evaluation method is a tool that uses a set of criteria, rather than a comparison group, for single-worksite assessments.	http://www.cdc.gov/nccdphp/dnpa/hwi/program_design/swat/index.htm
Reach, Effectiveness, Adoption, Implementation, and Maintenance (RE-AIM)	Workgroup members: David Dzewaltowski, Paul Estabrooks, Russ Glasgow, and Lisa Klesges; funded by RWJF	To implement and evaluate an explanatory framework for measuring intervention impact in its broadest sense. The framework is a systematic way for researchers and decision makers to evaluate health behavior interventions.	http://www.re-aim.org
Systematic Reviews			
Centre for Reviews and Dissemination	University of York and National Institute for Health Research	To undertake systematic reviews of research evidence on health and public health questions of national and international importance. The findings are widely disseminated and have impacted health care policy and practice in the United Kingdom and internationally.	http://www.york.ac.uk/inst/crd/index.htm

TABLE B-1 Continued

Project Name	Funder(s)/Investigator(s)	Purpose	Website
Cochrane Public Health Review Group	Cochrane Collaboration	To produce systematic reviews of population-level interventions; contribute to methodological developments in the synthesis of public health research; transfer knowledge and exchange initiatives; and develop strategies for primary research on and rigorous evaluations of population-level interventions (policy, programs, legislation, community interventions), including maintaining a specialized studies register.	http://www.ph.cochrane.org/en/index.html
Development of Evidence Standards for Child Health Promotion	Nemours Health and Prevention Services, The Commonwealth Fund, and the Agency for Healthcare Research and Quality	To propose and disseminate a set of clearly defined evidence standards appropriate for primary care that can be used to determine the effectiveness of preventive and developmental services for children and allow the advancement of the field through the proper development and analysis of scientific evidence.	http://www.commonwealthfund.org/Content/Grants/2006/Dec/Developing-New-Evidence-Standards-for-Child-Development-and-Health-Promotion-Services—Phase-2.aspx
Effective Public Health Practice Project (EPHPP)	Ontario Public Health Research, Education and Development Program	To conduct systematic reviews of the effectiveness of public health interventions and summarize recent high-quality reviews produced by others. The products of these reviews are disseminated as resources for evidence-based decision making in public health in Canada.	http://www.ephpp.ca/index.htm
Review of Environmental and Policy Interventions for Childhood Obesity Prevention (EPNP)	Transtria, LLC with RWJF and Washington University Public Health Institute	(1) To create a systematic annual review process for classifying environmental and policy interventions for childhood obesity prevention into levels of evidence. (2) To identify environmental and policy interventions in nutrition and physical activity with applicability in children aged 3-18 emerging from existing evidence-based review processes. (3) To review and summarize emerging evidence-based environmental and policy interventions in nutrition and physical activity with applicability in children aged 3-18 with respect to a range of research- and practice-based levels of evidence.	http://www.transtria.com/projects.php

continued

TABLE B-1 Continued

Project Name	Funder(s)/Investigator(s)	Purpose	Website
Uses of Evidence			
Collaboration of Community-based Obesity Prevention Sites (CO-OPS Collaboration)	Australian Government Department of Health and Ageing	To support community-based obesity prevention initiatives through a collaborative approach to promoting best practices and knowledge translation, and through the provision of networking opportunities, support, and advice.	http://www.co-ops.net.au/
Health Evidence Canada	National Collaborating Centre for Aboriginal Health (NCCAH), National Collaborating Centre for Environmental Health (NCCEH), National Collaborating Centre for Healthy Public Policy (NCCHPP), National Collaborating Centre for Infectious Diseases (NCCID), National Collaborating Centre for Methods and Tools (NCCMT), City of Hamilton Public Health Service Division	To facilitate the adoption and implementation of effective policies/programs/interventions at the local and regional public health decision-making levels across Canada, and to provide decision makers with tools that will enhance their quality assessment skills and to customize the content they receive to specified areas of interest.	http://health-evidence.ca
National Registry of Evidence-Based Programs and Practices	Substance Abuse and Mental Health Services Administration	To assist the public in identifying approaches to preventing and treating mental and/or substance use disorders that have been scientifically tested and that can be readily disseminated to the field.	http://www.nrepp.samhsa.gov/
Roundtable on Evidence-Based Medicine	Institute of Medicine	To convene leadership from key health care sectors to catalyze the collaborative work needed to drive improvements in the effectiveness and efficiency of medical care by transforming how evidence is developed and used in health care.	http://www.iom.edu/ebm
TREND—Transparent Reporting of Evaluations with Nonrandomized Designs	Various public health collaborators as an outgrowth of a CDC meeting	To develop standardized, transparent reporting standards for nonrandomized intervention research evaluations in public health–related fields.	http://www.cdc.gov/trendstatement/

C

Review of Existing Reviews on Obesity Prevention

The Committee on an Evidence Framework for Obesity Prevention Decision Making was tasked to provide an overview of the current nature of the evidence base, that is, the types of evidence that are available on the results of community, environmental, and policy-based obesity prevention initiatives. The narrative interpretation of the collection of reviews in this appendix (termed "appraisals" in this report) is provided in Chapter 3, and includes a description of current methods used to characterize the evidence base and a discussion of the challenges involved in applying traditional evidence hierarchies to population-based prevention efforts.

This overview was not meant to be carried out by an exhaustive literature review but rather through an examination of case studies and existing reviews. The search strategy involved searching major databases including Scopus, PubMed, and Web of Science, from each database's inception through November 2009, for relevant and published reviews of obesity prevention interventions. Variations of the terms "overweight," "obesity," "prevention," "physical activity," "nutrition," "review," "meta-analysis," "synthesis," "summary," and "recommendations" were used in the search process. Additional appraisals were identified by review of the reference lists of included appraisals and other relevant literature and by discussion with experts in the field. Appraisals that consider only studies of treatment interventions were excluded; some appraisals consider both treatment and prevention studies and were included in the committee's analysis. Publication date was not a search criterion, yet all of the appraisals in this appendix were published between 1997 and the present.

TABLE C-1

	Search Strategy				Inclusion/Exclusion Criteria
Study	Sources of Information[a]	Years Included	Type of Appraisal and Purpose	Initial No. Identified	Subject Demographics
American Dietetic Association Position Statement (2006)	(28) Manual searching of bibliographies from gathered articles and reviews	January 1982-January 2004	Systematic review To develop a position statement on individual-, family-, school-, and community-based interventions for pediatric overweight	Not reported	Healthy children and adolescents aged 2-18
Bluford et al. (2007)	(11), (14), (20), (27), (28)	1966-March 2005	Systematic review To identify and summarize effective evaluated interventions for preventing or treating obesity among preschool children aged 2-6	Not reported	Preschool-age children aged 2-6
Brown and Summerbell (2009)	(20), (28)	January 2006-September 2007	Systematic review To determine the effectiveness of school-based interventions focused on changing dietary intake and physical activity levels to prevent childhood obesity	1,553	School children aged 5-18 Excluded: studies that recruited children on the basis of weight; studies in children with critical illness or eating disorders

Types of Studies (e.g., RCTs, quasi-experimental, qualitative)	Outcome Measures (e.g., BMI, diet, physical activity, other behaviors, environment)	Length of Intervention and/or Follow-up	Other	No. After Applying Criteria
Included: any study design Excluded: meta-analyses and reviews (different selection criteria)	Measure of adiposity (BMI, body weight, skinfold thickness, percent body fat)	> 8 weeks (for secondary prevention trials) ≥ 6 months (for primary prevention trials) (not including duration of follow-up)	Included: English language Excluded: conducted in a developing country; not peer-reviewed; n < 30 in secondary prevention trials, n < 60 in primary prevention trials	Not reported
Included: domestic and international studies that used physical activity and nutritional strategies in interventions to prevent or treat overweight Excluded: case reports or series	Measure of weight status, BMI, or body fat	≥ 3 months study length	Studies were not excluded because of their aim (for example, did not have to include weight change), design, or duration of intervention or follow-up if ≥ 3 months	7 (2 treatment and 5 prevention)
RCTs or controlled clinical trials of lifestyle interventions Included: study designs that compared lifestyle interventions (defined in article) with usual care or with other active interventions	Included: same criteria used by National Institute for Health and Clinical Excellence (NICE) obesity guidance with one exception—included studies had to report a weight outcome, either as absolute or change, and could use BMI, BMI *z*-score, percentage of body fat, skin-fold thickness, and percentage of overweight	≥ 12 weeks study length	Included: minimum duration of 12 weeks; done in a school setting (studies were not included or excluded based on their aim)	38

continued

TABLE C-1 Continued

	Search Strategy				Inclusion/Exclusion Criteria
Study	Sources of Information[a]	Years Included	Type of Appraisal and Purpose	Initial No. Identified	Subject Demographics
Campbell et al. (2001)	(13), (14), (20), (28), (32), (33), (36) Expert opinion	January 1985-October 1999	Systematic review To assess the effectiveness of interventions designed to prevent obesity in childhood	1,043	Included: subjects under age 18 when the study began Excluded: pregnant women, those with eating disorders, and the critically ill
Connelly et al. (2007)	(14), (20), (27), (28) Manual searching of bibliographies from gathered articles and reviews	Up to April 30, 2006	Systematic review To present practice-relevant guidance on interventions aimed at reducing at least one measure of adiposity in child populations that do or do not contain overweight or obese children	Not reported	n ≥ 30; nonclinical child populations aged 0-18
Conroy et al. (2007)	(2), (12), (13), (14), (17), (19), (20), (23), (27), (28), (35), (38), (39), (43) Grey literature, excluding newspaper articles/editorials; manual searching of bibliographies from gathered articles and reviews; ancestry searches on key papers	1980-September 2005	Integrative review To examine successful Canadian nursing and health promotion intervention programs for childhood obesity prevention during gestation and infancy	2,028	Not reported

Types of Studies (e.g., RCTs, quasi-experimental, qualitative)	Outcome Measures (e.g., BMI, diet, physical activity, other behaviors, environment)	Length of Intervention and/or Follow-up	Other	No. After Applying Criteria
RCTs, non-RCTs with concurrent control groups	Included: had to report one or more outcomes (estimates of percent body fat, BMI, ponderal index, or skin-fold thickness) and report outcome data at baseline and at postintervention, or at baseline and change from baseline	Follow-up 1 year, later amended to include studies with an observation period of ≥ 3 months	Not reported	7
Included: RCTs, controlled trials of interventions to prevent overweight or obesity	Included: measure of index of adiposity	≥ 12 weeks study period	Not reported	28
Included: peer-reviewed and/or research-based, published articles, policy reports Excluded: newspaper articles, editorials, general reports, general women's health materials	Diet, BMI (adult), high birth weight > 4000 g, low birth weight < 2500 g but > 2000 g	Not reported	Canadian only; English and French language	6

continued

TABLE C-1 Continued

	Search Strategy				Inclusion/Exclusion Criteria
Study	Sources of Information[a]	Years Included	Type of Appraisal and Purpose	Initial No. Identified	Subject Demographics
DeMattia et al. (2007)	(27), (28) Updated search: (12), (14), (23), (27), (28) Manual searching of bibliographies from gathered articles and reviews; expert opinion	1966-June 2004 Updated search, 1966-February 2005	Systematic review To examine studies that used an intervention to reduce sedentary behavior (SB; defined in article) to control weight in order to answer the question: Is an emphasis on reducing SB (e.g., TV, video games) effective at changing behaviors and controlling weight in children and adolescents?	222	Children or adolescents, including at-risk or obese children/teens at outset of study
Doak et al. (2006)	(28) Personal contacts with researchers; Internet; manual searching of bibliographies from reviews; MEDLINE searches of authors with ongoing intervention studies	Through June 2003, with additional update in August 2005	Review To identify aspects of overweight and obesity prevention programs that are most likely to succeed if implemented on a large scale	102	Children and adolescents aged 6-19; targeting of obese children if drawn from a broad-based school population
Dobbins et al. (2001)	(14), (19), (20), (27), (28), (30), (38) Review of reference list of retrieved articles	1985-September 2000	Systematic review To summarize the evidence for the effectiveness of school-based interventions in promoting physical activity and fitness in children and adolescents	143	Children and adolescents aged 6-18

Types of Studies (e.g., RCTs, quasi-experimental, qualitative)	Outcome Measures (e.g., BMI, diet, physical activity, other behaviors, environment)	Length of Intervention and/or Follow-up	Other	No. After Applying Criteria
Included: RCTs, controlled clinical trials, comparative studies, multicenter studies that aimed to reduce SB in a natural setting Excluded: observational cross-over and cohort studies, studies of behavior in a controlled laboratory setting	Included: Measure of SB and weight (either BMI or % overweight)	Not reported	Validity criteria applied	12
Included: school-based studies seeking to alter diet and/or physical activity–related behaviors and having a broad public health base; studies that monitored and evaluated in a manner that has been documented (e.g., published) Excluded: studies with aims so narrow that an impact on obesity is unlikely; clinical settings focused on treatment of obese children; pilot studies	Included: anthropometric measurement of body weight or obesity, at baseline and follow-up	Not reported	Not reported	25
RCTs; studies with a control group and measures taken pre- and postintervention; prospective study designs	Outcomes must have been reported for children and adolescents		Included: intervention applicable to public health practice; intervention could be implemented, facilitated, or promoted by staff in local public health units in Ontario	19

continued

TABLE C-1 Continued

	Search Strategy				Inclusion/Exclusion Criteria
Study	Sources of Information[a]	Years Included	Type of Appraisal and Purpose	Initial No. Identified	Subject Demographics
Eden et al. (2002)	(7), (12), (13), (23), (28) Expert opinion; manual searching of bibliographies from pertinent articles	1994-March 2002	Systematic review To determine whether counseling adults in primary care settings improves and maintains physical activity levels	~1,000	Adults in general primary care populations

Types of Studies (e.g., RCTs, quasi-experimental, qualitative)	Outcome Measures (e.g., BMI, diet, physical activity, other behaviors, environment)	Length of Intervention and/or Follow-up	Other	No. After Applying Criteria
Included: controlled trials, case-control studies, observational studies	Included: behavioral outcomes (physical activity); if the study was of "good" or "fair" quality, according to criteria developed by the U.S. Preventive Services Task Force (USPSTF)	Not reported	The patient's primary care clinician performed some of the counseling intervention	8 (7 RCTs, 1 non-RCT)

continued

TABLE C-1 Continued

	Search Strategy				Inclusion/Exclusion Criteria
Study	Sources of Information[a]	Years Included	Type of Appraisal and Purpose	Initial No. Identified	Subject Demographics
Faith et al. (2007)	(3), (28), (32), (43), (33), (3) Manual searching of bibliographies from pertinent articles and chapters; expert opinion; grey literature (government and private foundation reports, excluded dissertations)	Earliest data indexed through November 2004	Narrative review To review critically the evidence addressing three questions that are pertinent to potential macro-level environmental influences on obesity management	Not reported	Not reported

	Types of Studies (e.g., RCTs, quasi-experimental, qualitative)	Outcome Measures (e.g., BMI, diet, physical activity, other behaviors, environment)	Length of Intervention and/or Follow-up	Other	No. After Applying Criteria
	For substudy 1: Included: focus group, observational study, quasi-experimental and experimental designs, RCTs Excluded: position or conceptual articles For substudy 2: Included: the macro-environmental factor must be studied in such a way that it was not confounded with education-based manipulations aimed at changing behavior through structured information dissemination Excluded: comprehensive multicomponent interventions targeting multiple behaviors (because the specific effects of manipulating individual components could not be disaggregated) For substudy 3: Excluded: studies testing the effects of restrictive parental feeding practices and strategies on child eating behavior and weight (these studies, being largely confined to the home, are less central to the topic of macro-environmental influences) For all substudies: Excluded: studies aimed primarily (or to a large degree) at achieving behavior change through information dissemination	Included: had to have at least one of the following: food acquisition or purchasing behavior, reported or weighed food intake, or body weight or composition	Not reported	Not reported	5 (substudy 1), 17 (substudy 2), 1 (substudy 3)

continued

TABLE C-1 Continued

	Search Strategy				Inclusion/Exclusion Criteria
Study	Sources of Information[a]	Years Included	Type of Appraisal and Purpose	Initial No. Identified	Subject Demographics
Flodmark et al. (2006)	(15), (28) Manual searching of bibliographies from relevant articles, journals, and reviews	Articles published until 2004 added to an earlier (2002) review by the Swedish Council on Technology Assessment in Health Care	Systematic literature review To review the scientific evidence for medical interventions aimed at preventing obesity during childhood and adolescence	136	Included: participants representing a normal group (school, or the population within a particular area), studies of high-risk groups Excluded: participants selected only for having the risk factor overweight/obesity (i.e., treatment)
Flynn et al. (2006)	(2), (3), (12), (13), (14), (17), (19), (20), (23), (24), (25), (27), (28), (29), (31), (38), (39) Manual searching of bibliographies from *International Journal of Obesity* (1987-2003), relevant reports; Internet searches of relevant domestic and foreign organization and society sites, Google; expert opinion	1982-2003	Synthesis research study To develop best practice recommendations based on a systematic approach to finding, selecting, and critically appraising programs addressing prevention and treatment of childhood obesity and related risk of chronic diseases	13,158 (library database hits)	Included: studies with a majority of participants under age 17 Excluded: those with compromised physical health; those with mental illness; programs with a primary focus on children of diabetic mothers, low-birth-weight/preterm babies, and pregnant adolescents

Types of Studies (e.g., RCTs, quasi-experimental, qualitative)	Outcome Measures (e.g., BMI, diet, physical activity, other behaviors, environment)	Length of Intervention and/or Follow-up	Other	No. After Applying Criteria
Study addresses prevention of obesity or overweight; must include a control group	Measure of BMI, skinfold thickness; % overweight/obesity	≥ 12 months with follow-up	Accepted Swedish, German, Norwegian, Danish, English, and French languages	24
Excluded: case studies; surgical interventions; programs associated with marketing products/services	Included: a measure of either outcome or process (or both) (1) an index of overweight/obesity: anthropometry, BMI, body fat distribution, growth rates (2) risk factors for obesity (e.g., activity levels) (3) chronic diseases (or risk factors) associated with overweight/obesity (e.g., type 2 diabetes or blood pressure) (4) adverse effects (5) process indicators (e.g., no. invited vs. no. participating, information evaluating how the program proceeded, or description of participant groups)	Not reported	Excluded: foreign-language articles without an English abstract	158

continued

TABLE C-1 Continued

	Search Strategy				Inclusion/Exclusion Criteria
Study	Sources of Information[a]	Years Included	Type of Appraisal and Purpose	Initial No. Identified	Subject Demographics
Galani and Schneider (2007)	(14), (20), (28) Manual searching of bibliographies from included articles, relevant books, individual-based studies of type 2 diabetes mellitus and cardiovascular disease	1995-2005 (because of advances in research toward conducting higher-quality studies and better reporting)	Systematic literature review with meta-analysis To assess the mid- to long-term effectiveness of lifestyle interventions in the prevention and treatment of obesity (defined in article)	273 (for prevention)	Children ≥ 18 years old
Glenny et al. (1997)	(20), (28), (32), Bath Information and Data Services (BIDS) (depending on author's subscription) Manual searching of bibliographies from gathered articles and reviews; expert opinion	Not reported	Systematic review To determine the effectiveness of interventions designed to prevent and treat obesity and maintain weight loss	97 for both prevention and treatment	Overweight or obese children and adults
Gonzalez-Suarez et al. (2009)	(1), (4), (6), (9), (14), (15), (16), (20), (22), (28), (34), (35), (41), Ovid (depending on author's subscription) Manual searching of bibliographies from relevant studies	1995-2007 (English only)	Meta-analysis To evaluate the effectiveness of school-based programs in the prevention and management of childhood obesity	41	Children and adolescents (schoolchildren) Normal BMI only

Types of Studies (e.g., RCTs, quasi-experimental, qualitative)	Outcome Measures (e.g., BMI, diet, physical activity, other behaviors, environment)	Length of Intervention and/or Follow-up	Other	No. After Applying Criteria
Included: RCTs of lifestyle interventions (defined in article)	Included: body weight, BMI, waist circumference, systolic/diastolic blood pressure, total cholesterol, LDL, HDL, triglycerides, glucose control (2-hour plasma glucose, fasting plasma glucose, or 3-month average blood glucose level [A1c])	≥ 12 months observation period	Included: prevention and treatment studies	13 (for prevention)
RCTs (for treatment of obesity and maintenance of weight loss); RCTs and nonrandomized studies with a concurrent control group (for prevention of obesity)	Postintervention changes relevant to baseline in one of the following: weight, fat content, fat distribution	≥ 12 months (for intervention only, or intervention + follow-up)	Not reported	4 (for prevention)
Included: RCTs and clinical controlled trials in which odds ratios or standardized mean differences and their 95% confidence intervals were reported or could be calculated	BMI, waist girth, percentage body fat, triceps skinfold, waist-to-hip ratio	Not limited	Seeking interventions that increase physical activity, improve dietary behaviors, modify poor exercise or dietary behaviors, or a combination of strategies	19

continued

TABLE C-1 Continued

	Search Strategy				Inclusion/Exclusion Criteria
Study	Sources of Information[a]	Years Included	Type of Appraisal and Purpose	Initial No. Identified	Subject Demographics
Hardeman et al. (2000)	(15), (16), (19), (20), (23), (28), (32), (36) Manual searching of bibliographies from gathered articles Queried authors of included papers	20: 1980-present; 28: 1966-present; 32: 1974-present	Systematic review To identify and review published studies on interventions aimed at the prevention of weight gain	Not reported	Participants not selected by weight or age Excluded: interventions in specific subgroups (e.g., those stopping smoking)
Harris et al. (2009)	(13), (14), (20), (28) Manual searching of bibliographies from relevant journals and articles	As early as 1966 for some databases, up to September 2008 for all databases	Systematic review and meta-analysis To determine whether school-based physical activity interventions improve children's body composition, as measured by BMI	398	Children aged 5-18
Hudson (2008)	(14), (27), (28)	1997-2007	Integrative review To explore attitudes, perceptions, and lifestyle behaviors that influence obesity risk in African American children	Not reported	Included: African American children aged 2-18 or their parents/caretakers Excluded: studies targeting subjects with BMI ≥ 85%

Types of Studies (e.g., RCTs, quasi-experimental, qualitative)	Outcome Measures (e.g., BMI, diet, physical activity, other behaviors, environment)	Length of Intervention and/or Follow-up	Other	No. After Applying Criteria
All study designs included	Data extracted based on: behaviors targeted for change; psychological model; behavior change methods and modes of delivery; methodological quality; characteristics of participants; outcomes related to body weight; self-reported diet and physical activity	Any duration	Included: published reports of interventions with the primary aim of preventing weight gain Excluded: if weight control was part of a multifactorial intervention aimed primarily at preventing a specific disease; if weight loss was primary aim	11
Included: RCTs, non-RCTs with an appropriate control group	Included: objective data for BMI from before and after the intervention (in both intervention and control groups)	≥ 6 months	Included: school-based exercise or physical activity interventions	18
Included: qualitative, experimental, and nonexperimental designs Excluded: prevalence studies	Included: BMI for age/gender or BMI z-score as primary outcome Excluded: studies not reporting outcomes by racial/ethnic characteristics	Not a criterion	Included: English, published in peer-reviewed journal Excluded: non-U.S. studies	28

continued

TABLE C-1 Continued

	Search Strategy				Inclusion/Exclusion Criteria
Study	Sources of Information[a]	Years Included	Type of Appraisal and Purpose	Initial No. Identified	Subject Demographics
Jago and Baranowski (2004)	(28) Manual searching of bibliographies from gathered articles and reviews; contacting of leading authors in the field for copies of articles "in press"	1970-2002	Systematic review To study noncurricular approaches to increasing physical activity in children and adolescents	Not reported	Children aged 5-18
Kahn et al. (2002)	(21), (27), (28), (37), (38), (39), (40) Manual searching of bibliographies from gathered reviews; expert opinion	1980-2000	Systematic review and meta-analysis To evaluate the effectiveness of various approaches to increasing physical activity: informational, behavioral and social, and environmental and policy approaches	6,238	Included: studies that compared outcomes among groups Excluded: studies that used physical activity solely as a therapeutic intervention; training studies
Kamath et al. (2008)	(13), (14), (19), (20), (27), (28), (30), (33), (36) Manual searching of bibliographies from gathered articles and reviews; expert opinion	Inception of the databases until February 2006	Systematic review and meta-analysis To summarize evidence on the efficacy of interventions aimed at changing lifestyle behaviors to prevent obesity	1,162	Included: children and adolescents aged 2-18 Excluded: prevention reviews focused exclusively on obese children; studies with participants with eating disorders, underlying diagnosis predisposing to obesity and/or inhibiting normal growth; adults or patients with type 1 diabetes

Types of Studies (e.g., RCTs, quasi-experimental, qualitative)	Outcome Measures (e.g., BMI, diet, physical activity, other behaviors, environment)	Length of Intervention and/or Follow-up	Other	No. After Applying Criteria
Excluded: case reports, unpublished reports, dissertations	Included: any increase in the level of physical activity; objective and self-report assessments of physical activity; interventions that measured physical activity before and after the intervention	Not reported	Included: English language; effect of a noncurricular physical activity intervention was evaluated	9
Included: interventions selected for evaluation, not guidelines or reviews; study design with a concurrent or before-and-after comparison	Included: aerobic capacity as an indicator of physical activity behavior	Not reported	Used *Guide to Community Preventive Services*' methods (CDC, 2005); English language only	94
Included: RCTs only Excluded: treatment articles	Included: BMI; behavioral endpoints of dietary or physical activity changes (e.g., accelerometer); self-report or reported by a family member or responsible adult, research or health care personnel	Used longest period of follow-up where data were available and not excessive loss to follow-up (> 20%)	Review conducted in accordance with Cochrane Collaboration methods and Quality of Reporting of Meta-analysis of Randomized Trials (QUOROM) standards (Moher et al., 1999) Excluded: foreign language studies not yet translated	34, BMI; 29, behavioral endpoints (usable data for meta-analysis)

continued

TABLE C-1 Continued

	Search Strategy				Inclusion/Exclusion Criteria
Study	Sources of Information[a]	Years Included	Type of Appraisal and Purpose	Initial No. Identified	Subject Demographics
Katz et al. (2005)	(15), (20), (23), (27), (28) Manual searching of bibliographies from gathered reviews and reports	1966-2001	Systematic review To identify effective strategies for weight control in school or work settings involving nutrition, physical activity, combinations of these, and other behavioral interventions	44 (school setting), 35 (worksite)	School-based interventions up to age 18, and workplace-based interventions (working-age adults)
Katz et al. (2008)	(14), (20), (23), (27), (28), Ovid (depending on database access) Manual searching in (15) and other meta-analyses	(20), (23), (27), (28): 1966-February 2000 (14), (27), (28), Ovid: February 2000-October 2004	Systematic review and meta-analysis To evaluate the effectiveness of school-based strategies for obesity prevention and control	64	Included: studies targeting children aged 3-18 in a school setting
Khan et al. (2009)	(28) Expert opinion	January 1, 2005-July 3, 2007	Systematic review To identify and recommend a set of strategies and associated measurements that communities and local governments can use to plan and monitor environmental and policy-level changes for obesity prevention	270	Included: policy and environmental-level strategies

Types of Studies (e.g., RCTs, quasi-experimental, qualitative)	Outcome Measures (e.g., BMI, diet, physical activity, other behaviors, environment)	Length of Intervention and/or Follow-up	Other	No. After Applying Criteria
Included: used *Guide to Community Preventive Services* (RCTs); control measurement between or within groups	Included: Weight loss of ≥ 4 pounds, intermediate outcomes: change in diet or physical activity levels; common weight-related measures used (BMI, body weight, anthropometic measures); in the school setting, meaningful changes in weight based on intervention goal and population characteristics; among adults, 4-pound minimum weight loss	≥ 6 months after intervention start	Used methods of the *Guide to Community Preventive Services* (CDC, 2005) for conducting systematic reviews to assess study design, validity, quality limitations; English only	10 (school setting) 20 (worksite)
Included: studies that contained a control measurement (pre/post, control group) Excluded: studies with "poor methodological quality"; 11 studies because of data usability limitations	Included: common weight outcomes (BMI, body weight)	Follow-up at least 6 months from the beginning of the intervention	Used *Guide to Community Preventive Services* (CDC, 2005) to abstract data and assess methodological quality; English only	8 (combination of nutrition and physical activity interventions)
Not reported	Not reported; used criteria to rate measurements: utility, construct validity, feasibility	Not reported	Common Community Measures for Obesity Prevention Project Team (CDC, experts, consultants); used criteria to rate strategies: reach, mutability, transferability, effect size, sustainability of health impact	179

continued

TABLE C-1 Continued

	Search Strategy				Inclusion/Exclusion Criteria
Study	Sources of Information[a]	Years Included	Type of Appraisal and Purpose	Initial No. Identified	Subject Demographics
Kropski et al. (2008)	(8), (18), (28) Manual searching of bibliographies from gathered articles; expert opinion	January 1, 1990-December 31, 2005	Systematic review To examine the effectiveness of school-based programs for reducing childhood overweight or obesity	Unknown	Children (school-based programs)
Lemmens et al. (2008)	(15), (20), (27), (28), (43), Ovid (depending on author's subscription)	January 1996-June 2006	Systematic review To examine the efficacy of obesity prevention interventions in adults, with a focus on dietary intake, physical activity, or a combination of these	13,198	Included: nonobese adults (> age 18) with an average BMI at the starting point < 30 or overweight/not yet obese (25-30) Excluded: those with BMI higher than 30; weight maintenance interventions; institutionalized adults; those with chronic conditions; drug interventions; health care providers, preoperative settings; bariatric surgery
Matson-Koffman et al. (2005)	(11), (19), (22), (27), (28), (43) Reports of U.S. Department of Agriculture, U.S. Department of Transportation	1970-October 2003	Narrative review To assess policy and environmental interventions aimed at increasing physical activity or improving nutrition	Not reported	Included: all demographics

Types of Studies (e.g., RCTs, quasi-experimental, qualitative)	Outcome Measures (e.g., BMI, diet, physical activity, other behaviors, environment)	Length of Intervention and/or Follow-up	Other	No. After Applying Criteria
Included: experimental or quasi-experimental designs	Included: BMI, a measure of body fat or obesity/overweight prevalence	Evaluation ≥ 6 months post baseline	Used Grading of Recommendations Assessment, Development, and Evaluation (GRADE) criteria for analyzing evidence (Atkins et al., 2004, 2005)	14
Included: RCTs, nonrandomized trials, and interrupted time series designs; with control group	Included: BMI or body weight change, self-report or measured	Minimum 3-month follow-up	Used *Cochrane Collaboration Handbook* for systematic review protocol	9 (5 long-term, at least 1 year; 4 short-term, 3 months to 1 year)
Not reported	Included: behavioral, psychological, or organizational change outcomes	Not reported	Excluded: studies that did not provide a description of the intervention or those that reported on determinants research (individual-level only) or media-only campaigns; studies that linked the built environment to physical activity (in another review)	65

continued

TABLE C-1 Continued

	Search Strategy				Inclusion/Exclusion Criteria
Study	Sources of Information[a]	Years Included	Type of Appraisal and Purpose	Initial No. Identified	Subject Demographics
Mayer (2009)	(14), (19), (27), (28), (43) Internet (Google)	2001-January 2008	Integrative review To assess community-based obesity prevention interventions focused on changing the nutrition environment for children	356	Included: those aged 2-21 years Excluded: adults (> 21 years)
Micucci et al. (2002)	(14), (15), (19), (20), (27), (28), (36) Expert opinion to find unpublished material	1990-January 2002	Review of reviews using a chronic disease model (i.e., risk factors for type 2 diabetes) To assess the effectiveness of school-based interventions in primary prevention of obesity and promotion of physical activity and nutrition	1,448 (195 with potential relevance)	Included: school-aged youth in a school setting
Pignone et al. (2003)	(28) Manual searching of bibliographies from gathered articles and reviews; expert opinion	1966-December 2001	Systematic review To examine the effectiveness of counseling to promote healthful diet among patients in primary care settings	129 (74 reported on counseling on dietary behavior)	Included: studies with patients with known risk factors for chronic disease Excluded: studies with patients that had previously diagnosed chronic illness or required special diets; obese or overweight patient studies for the purpose of weight management

Types of Studies (e.g., RCTs, quasi-experimental, qualitative)	Outcome Measures (e.g., BMI, diet, physical activity, other behaviors, environment)	Length of Intervention and/or Follow-up	Other	No. After Applying Criteria
Included: RCTs, pre–post trials, longitudinal, follow-up studies Excluded: systematic reviews, descriptive articles, treatment-focused studies	Did not limit these criteria, but reported what they found	Did not limit this criterion, but reported what they found	English language	12
Not specified, but followed *Cochrane Collaboration Handbook*	Included: physiological indicators (body fat, metabolic indices, change in physical activity, healthful eating/nutrition)	Not reported	Followed the guidelines of the *Cochrane Collaboration Handbook*; quality ratings used; only strong and moderate reviews included	12 (7 strong, 5 moderate)
Included: RCTs Excluded: studies with dropout rates of > 50% or with presence of a disease	Included: dietary behavior; net change in consumption Excluded: studies with only biochemical or anthropomorphic endpoints	At least 3 months' duration	Excluded: studies that controlled the diet externally	21

continued

TABLE C-1 Continued

	Search Strategy				Inclusion/Exclusion Criteria
Study	Sources of Information[a]	Years Included	Type of Appraisal and Purpose	Initial No. Identified	Subject Demographics
Rees et al. (2006)	(15), (19), (27), (36)	Coverage of database until 2001	Systematic review To examine the barriers to and facilitators of physical activity among young people, focused on examining community- and society-level interventions	7,048 (90 studies reported on physical activity)	Included: youth aged 11-16
Reilly and McDowell (2003)	(14), (15), (20), (23), (28) Manual searching of bibliographies from gathered systematic reviews and individual intervention articles; Internet	June 2000-May 2002	Systematic review To examine interventions for prevention and treatment of childhood obesity targeting physical activity (included separate analyses of treatment and prevention)	Not reported	Included: children and adolescents; nonclinical groups of subjects

Types of Studies (e.g., RCTs, quasi-experimental, qualitative)	Outcome Measures (e.g., BMI, diet, physical activity, other behaviors, environment)	Length of Intervention and/or Follow-up	Other	No. After Applying Criteria
Included: trials and studies that used a comparison group design and reported both pre- and posttest data; also outcome or process evaluation, systematic review or study that did not involve an intervention, and nonintervention studies that studied young people's definitions of ideas about physical activity or factors that influence their own or other young people's physical activity	Measured behavior or health of young people	Not reported	Included: interventions that aimed to make a change at the level of the community or society; quality of trials assessed using the Evidence for Policy and Practice Information and Coordinating Centre (EPPI-Centre) review guidelines (Peersman et al., 1997); quality of people's views studies as described by Harden et al. (2004)	12 trials, 16 studies of young people's views
Included: RCTs	Included: body weight, BMI, body composition	At least 12 months	Used *Checklist for Appraising the Quality of Controlled Trials* (Scottish Intercollegiate Guidelines Network, 2002) for quality appraisal	3

continued

TABLE C-1 Continued

	Search Strategy				Inclusion/Exclusion Criteria
Study	Sources of Information[a]	Years Included	Type of Appraisal and Purpose	Initial No. Identified	Subject Demographics
Renzaho et al. (2010)	(14), (26), (27), (28), (34), (38), (42) Internet (Google Scholar)	Until September 2008	Systematic review To study the effectiveness of prevention programs for obesity and chronic diseases among immigrants in developed countries	599	Included: studies targeting immigrants Excluded: studies focused on nonimmigrants; specifically, at least 40% of multiethnic samples needed to be non–African American migrants
Salmon et al. (2007)	(14), (15), (26), (27), (28), (34), (37), (39), (43), Ovid (depending on access)	January 1985-June 2006	Narrative review To study the effectiveness of interventions in studies that report physical activity outcomes in children and adolescents	Not reported	Included: children aged 4-12, adolescents aged 13-19
Sharma (2007)	(14), (19), (28)	1999-2005	Narrative review To review international school-based interventions for preventing childhood obesity	21	Included: general population of children aged 3-18, in a school setting Excluded: studies that focused solely on overweight/obese children or adolescents
Small et al. (2007)	(14), (27), (28)	Not reported	Narrative review To conduct a critical appraisal of current evidence on prevention and early treatment of overweight and obesity in young children	12	Included: children aged 4-7

Types of Studies (e.g., RCTs, quasi-experimental, qualitative)	Outcome Measures (e.g., BMI, diet, physical activity, other behaviors, environment)	Length of Intervention and/or Follow-up	Other	No. After Applying Criteria
Included: RCTs; quasi-randomized, controlled before-and-after studies	Included: BMI (weight, waist and hip measurements), lean tissue mass, body fat; nutritional targets for patients with diabetes; self-reported nutrition adherence for diabetes self-management; diabetes knowledge; self-efficacy and self-management; dietary fat consumption; glycemic control; cardiovascular risk factors	Not reported	Included: intervention group needed at least 40 subjects Excluded: "poor-quality" studies; medical and/or drug interventions; studies that did not define the migrant group or intervention or clearly outline the methodological approach	13
Included: RCTs, group randomized trials, quasi-experimental study designs; postintervention, assessment-only designs or no control (decided on a case-by-case basis)	Included: physical activity outcomes (survey, observational, and objective)	Not reported	Included: sample size of more than 16 Excluded: obesity treatment	99
Excluded: studies published in non–peer reviewed journals	Not reported, but ended up with studies that targeted nutrition and physical activity behaviors	Not reported, but a majority were 1 academic year long	Included: studies outside of U.S. only English language	12
Included: RCTs, systematic reviews, meta-analyses	Not reported	Not reported	Looking to see what they could find to critically appraise the current evidence	6 prevention; 6 treatment

continued

TABLE C-1 Continued

	Search Strategy				Inclusion/Exclusion Criteria
Study	Sources of Information[a]	Years Included	Type of Appraisal and Purpose	Initial No. Identified	Subject Demographics
Stice et al. (2006)	(14), (27), (28), (30) Manual searching of tables of contents of pertinent journals (*Preventive Medicine*, *Journal of Pediatrics*, *Health Education Quarterly*), narrative reviews, gathered articles; expert opinion for unpublished articles	1980-2005 (October)	Systematic review and meta-analysis To investigate obesity prevention programs and their effects; to identify features of the program associated with these effects	Not reported	Included: children and youth (up to age 22)
Summerbell et al. (2005)	(13), (14), (20), (27), (28) Manual searching of bibliographies from gathered articles and reviews; Internet; expert opinion	1990-February 2005	Systematic review To assess the effectiveness of interventions designed to prevent obesity in childhood through diet, physical activity, and/or lifestyle and social support	828	Included: children below age 18

Types of Studies (e.g., RCTs, quasi-experimental, qualitative)	Outcome Measures (e.g., BMI, diet, physical activity, other behaviors, environment)	Length of Intervention and/or Follow-up	Other	No. After Applying Criteria
Included: RCTs, evaluations of obesity prevention programs, trials; needed relevant comparison group, logistic regression or survival models	Included: proxy measure of body fat (BMI, skinfold thickness); effect sizes for weight gain (not self-report)	Not reported	Excluded treatment	46 trials (61 programs)
Included: controlled clinical trials (with and without randomization)	Included: primary outcomes: weight and height, percent fat content, BMI, Ponderal index, skinfold thickness; secondary outcomes: activity levels, dietary intake, change in knowledge, environmental change, stakeholders' views of findings, measures of self-esteem, etc., harm associated with process or outcomes, cost-effectiveness	Minimum duration of 12 weeks (short-term)	Used *Cochrane Collaboration Handbook*, health promotion theory Excluded: treatment	22 (10 long-term, 12 short-term)

continued

TABLE C-1 Continued

	Search Strategy				Inclusion/Exclusion Criteria
Study	Sources of Information[a]	Years Included	Type of Appraisal and Purpose	Initial No. Identified	Subject Demographics
Thomas (2006)	See above	1985-2003	See above	See above	Included: studies with elementary or secondary school-aged children, school component
Thomas et al. (2004)	(10), (12), (13), (14), (19), (20), (27), (28), (38), (39) Manual searching of 24 relevant peer-reviewed journals (June 1998-June 2003)	1984-2003	Systematic review To examine the effectiveness of interventions for promoting healthy weight, preventing overweight/obesity, and increasing physical activity among school-aged children and youth	More than 1,100	Children and adolescents aged 6-18
Timperio et al. (2004)	(27), (28), (39)	1999-September 2003	Review (update) To study the effectiveness of intervention strategies to promote physical activity among children, adolescents, and young adults	Not reported	Included: children, adolescents, young adults

Types of Studies (e.g., RCTs, quasi-experimental, qualitative)	Outcome Measures (e.g., BMI, diet, physical activity, other behaviors, environment)	Length of Intervention and/or Follow-up	Other	No. After Applying Criteria
Included: studies with a comparison group; RCTs and other study designs	Included: self-reported dietary intake, BMI, skinfold thickness, moderate to vigorous physical activity (MVPA) changes	See above	Used a recent review to critique the results of primary studies related to the effectiveness of physical activity and nutrition programs for obesity prevention in children and youth	57
Included: RCTs, cohort studies (prospective, with a control group)	Included: changes in total caloric intake and percent calories from fat, fruit, and vegetable intake; changes in body fat; changes in perception of body image; self-reported changes in duration, rate, and frequency of physical activity; changes in lung capacity (VO^2 max); self-reported changes in inactivity; BMI Excluded: studies that measured only changes in vitamin or fiber intake, changes in knowledge and attitudes re physical activity and nutrition	Not reported	Excluded: studies related to therapeutic nutrition and physical activity programs and those that included children with known illnesses or obese/overweight children; quality assessed using Mulrow et al. (1997) and Jadad et al. (1996)	425
Included: RCTs	Not reported	Not reported	Review of new evidence (update)	31

continued

TABLE C-1 Continued

	Search Strategy				Inclusion/Exclusion Criteria
Study	Sources of Information[a]	Years Included	Type of Appraisal and Purpose	Initial No. Identified	Subject Demographics
Van Sluijs et al. (2007)	(20), (28), (32), (35), (39) Manual searching of bibliographies from gathered articles	Up to December 2006	Systematic review To study the effectiveness of interventions aimed at promoting physical activity in children and adolescents	3,045	Included: children and adolescents (< 18 years) Excluded: participants with a specific disease or health problem
Wareham et al. (2005)	(28), (32)	2000-2005	Critical review To study the effectiveness of physical activity interventions for obesity prevention	Not reported	Included: adults and children
Wofford (2008)	(15), (28) Manual searching of bibliographies from gathered articles; Joanna Briggs Institute database (depending on author's subscription)	Not reported	Systematic review To review evidence related to obesity prevention in young children	> 5,000, then limited by date (2004-2008), but included seminal articles, articles on preschool-aged children and relevance to prevention of obesity	Included: Eventually limited search to preschoolers (prevention)

Types of Studies (e.g., RCTs, quasi-experimental, qualitative)	Outcome Measures (e.g., BMI, diet, physical activity, other behaviors, environment)	Length of Intervention and/or Follow-up	Other	No. After Applying Criteria
Included: controlled trials (with or without randomization); control group (non–physical activity intervention)	Included: outcomes related to physical activity (self-reported or objectively measured)	Not reported	Used 10-item quality assessment scale for methodological quality (internal validity)	57: 33 children, 24 adolescents
Included: observational studies (longitudinal, prospective cohort) Excluded: cross-sectional studies	Included: physical activity or body composition change	Not reported		30: 14 adults, 16 children
Not reported	Not reported	Not reported	The 41 selected for review were chosen because of their relevance to the prevention of overweight and obesity in the preschool population and the “stronger” quality of evidence they presented English language	41

continued

TABLE C-1 Continued

	Search Strategy				Inclusion/Exclusion Criteria
Study	Sources of Information[a]	Years Included	Type of Appraisal and Purpose	Initial No. Identified	Subject Demographics
Zenzen and Kridli (2009)	(14), (27), (28) Manual searching of bibliographies from gathered articles	2000-2007	Integrative review To study the degree of variability in the methodological approaches and theoretical frameworks of school-based obesity prevention programs, focused on the following interventions: dietary, physical activity, and healthy lifestyle education; parental involvement	Not reported	Included: children aged 4-18 (kindergarten through high school)

NOTES: BMI = body mass index; CDC = Centers for Disease Control and Prevention; N/A = not available; RCT = randomized controlled trial.

[a]Electronic databases are denoted numerically as outlined below:
(1) Academic Search Elite
(2) ACP Journal Club (American College of Physicians Evidence Based Medicine)
(3) AGRICOLA
(4) AMED (Allied and Complementary Medicine Database)
(5) ARC Service (Agricultural Research Council)
(6) AustHealth
(7) Best Evidence Database
(8) Biological Abstracts
(9) BioMed Central
(10) BIOSIS Previews
(11) Chronic Disease Prevention File (National Center for Chronic Disease Prevention and Health Promotion)
(12) CDSR (Cochrane Database of Systematic Reviews)
(13) CENTRAL/CCTR (Cochrane Central Register of Controlled Trials)
(14) CINAHL (Cumulative Index of Nursing and Allied Health Literature)
(15) Cochrane Library
(16) Current Contents—Life Sciences
(17) DARE (Database of Abstracts of Reviews of Effects)
(18) Education Abstracts
(19) ERIC (Education Resources Information Centre)
(20) EMBASE (Excerpta Medica Database)
(21) Environline

Types of Studies (e.g., RCTs, quasi-experimental, qualitative)	Outcome Measures (e.g., BMI, diet, physical activity, other behaviors, environment)	Length of Intervention and/or Follow-up	Other	No. After Applying Criteria
Included: experimental and nonexperimental studies	Not reported	Not reported	Used conceptual framework of Cooper (1984) to guide the systematic review; used Stetler et al. (1998) for assessing quality Included: school-based curriculum programs for obesity prevention; studies needed to manipulate at least one of the variables of dietary habits, physical activity, healthful lifestyle education, and parental involvement; English language	16

(22) Health Source: Nursing/Academic Edition
(23) HealthSTAR (Health Services Technology, Administration, and Research)
(24) PAIS (Public Affairs Information Service) International
(25) ProceedingsFirst
(26) PsycARTICLES
(27) PsycINFO (Psychological abstracts)
(28) PubMed or MEDLINE
(29) Population Index
(30) Proquest Digital Dissertation Abstracts International
(31) Proquest Digital Dissertations and Theses
(32) PsychLit—Silver Platter
(33) Science Citation Index (Web of Science)
(34) ScienceDirect
(35) SCOPUS
(36) Social Science Citation Index (Web of Science)
(37) Social SciSearch
(38) Sociological Abstracts
(39) SPORTDiscus
(40) TRIS (Transportation Research Information Services)
(41) TRIP (Turning Research into Practice)
(42) Web of Knowledge
(43) Web of Science

REFERENCES

American Dietetic Association. 2006. Position of the American Dietetic Association: Individual-, family-, school-, and community-based interventions for pediatric overweight. *Journal of the American Dietetic Association* 106(6):925-945.

Atkins, D., D. Best, P. A. Briss, M. Eccles, Y. Falck-Ytter, S. Flottorp, G. H. Guyatt, R. T. Harbour, M. C. Haugh, D. Henry, S. Hill, R. Jaeschke, G. Leng, A. Liberati, N. Magrini, J. Mason, P. Middleton, J. Mrukowicz, D. O'Connell, A. D. Oxman, B. Phillips, H. J. Schunemann, T. T. Edejer, H. Varonen, G. E. Vist, J. W. Williams, Jr., and S. Zaza. 2004. Grading quality of evidence and strength of recommendations. *British Medical Journal* 328(7454):1490.

Atkins, D., P. A. Briss, M. Eccles, S. Flottorp, G. H. Guyatt, R. T. Harbour, S. Hill, R. Jaeschke, A. Liberati, N. Magrini, J. Mason, D. O'Connell, A. D. Oxman, B. Phillips, H. Schunemann, T. T. Edejer, G. E. Vist, and J. W. Williams, Jr. 2005. Systems for grading the quality of evidence and the strength of recommendations II: Pilot study of a new system. *BMC Health Service Research* 5(1):25.

Bluford, D. A. A., B. Sherry, and K. S. Scanlon. 2007. Interventions to prevent or treat obesity in preschool children: A review of evaluated programs. *Obesity* 15(6):1356-1372.

Brown, T., and C. Summerbell. 2009. Systematic review of school-based interventions that focus on changing dietary intake and physical activity levels to prevent childhood obesity: An update to the obesity guidance produced by the National Institute for Health and Clinical Excellence. *Obesity Reviews* 10(1):110-141.

Campbell, K., E. Waters, S. O'Meara, and C. Summerbell. 2001. Interventions for preventing obesity in childhood: A systematic review. *Obesity Reviews* 2(3):149-157.

CDC (Centers for Disease Control and Prevention). 2005. *The guide to community preventive services: What works to promote health?* New York: Oxford University Press.

Connelly, J. B., M. J. Duaso, and G. Butler. 2007. A systematic review of controlled trials of interventions to prevent childhood obesity and overweight: A realistic synthesis of the evidence. *Public Health* 121(7):510-517.

Conroy, S., R. Ellis, C. Murray, and J. Chaw-Kant. 2007. An integrative review of Canadian childhood obesity prevention programmes. *Obesity Reviews* 8(1):61-67.

Cooper, H. 1984. *The integrative research review: A systematic approach*. Beverly Hills: Sage.

DeMattia, L., L. Lemont, and L. Meurer. 2007. Do interventions to limit sedentary behaviours change behaviour and reduce childhood obesity? A critical review of the literature. *Obesity Reviews* 8(1):69-81.

Doak, C. M., T. L. S. Visscher, C. M. Renders, and J. C. Seidell. 2006. The prevention of overweight and obesity in children and adolescents: A review of interventions and programmes. *Obesity Reviews* 7(1):111-136.

Dobbins, M., D. Lockett, I. Michel, J. Beyers, L. Feldman, J. Vohra, and S. Micucci. 2001. *The effectiveness of school-based interventions in promoting physical activity and fitness among children and youth: A systematic review.* Hamilton, ON: Effective Public Health Practice Project.

Eden, K. B., C. T. Orleans, C. D. Mulrow, N. J. Pender, and S. M. Teutsch. 2002. Does counseling by clinicians improve physical activity? A summary of the evidence for the U.S. Preventive Services Task Force. *Annals of Internal Medicine* 137(3):208-215.

Faith, M. S., K. R. Fontaine, M. L. Baskin, and D. B. Allison. 2007. Toward the reduction of population obesity: Macrolevel environmental approaches to the problems of food, eating, and obesity. *Psychological Bulletin* 133(2):205-226.

Flodmark, C. E., C. Marcus, and M. Britton. 2006. Interventions to prevent obesity in children and adolescents: A systematic literature review. *International Journal of Obesity* 30(4):579-589.

Flynn, M. A. T., D. A. McNeil, B. Maloff, D. Mutasingwa, M. Wu, C. Ford, and S. C. Tough. 2006. Reducing obesity and related chronic disease risk in children and youth: A synthesis of evidence with 'best practice' recommendations. *Obesity Reviews* 7(Supplement 1):7-66.

Galani, C., and H. Schneider. 2007. Prevention and treatment of obesity with lifestyle interventions: Review and meta-analysis. *International Journal of Public Health* 52(6):348-359.

Glenny, A. M., S. O'Meara, A. Melville, T. A. Sheldon, and C. Wilson. 1997. The treatment and prevention of obesity: A systematic review of the literature. *International Journal of Obesity* 21(9):715-737.

Gonzalez-Suarez, C., A. Worley, K. Grimmer-Somers, and V. Dones. 2009. School-based interventions on childhood obesity: A meta-analysis. *American Journal of Preventive Medicine* 37(5):418-427.

Hardeman, W., S. Griffin, M. Johnston, A. L. Kinmonth, and N. J. Wareham. 2000. Interventions to prevent weight gain: A systematic review of psychological models and behaviour change methods. *International Journal of Obesity* 24(2):131-143.

Harden, A., J. Garcia, and S. Oliver. 2004. Applying systematic review methods to studies of people's views: An example from public health research. *Journal of Epidemiology & Community Health* 58:794-800.

Harris, K. C., L. K. Kuramoto, M. Schulzer, and J. E. Retallack. 2009. Effect of school-based physical activity interventions on body mass index in children: A meta-analysis. *Canadian Medical Association Journal* 180(7):719-726.

Hudson, C. E. 2008. An integrative review of obesity prevention in African American children. *Issues in Comprehensive Pediatric Nursing* 31(4):147-170.

Jadad, A., R. A. Moore, M. D. Carroll, C. Jenkinson, D. J. Reynolds, and D. Gavaghan. 1996. Assessing the quality of reports of randomized clinical trials: Is blinding necessary. *Controlled Clinical Trials* 17(1):1-12.

Jago, R., and T. Baranowski. 2004. Non-curricular approaches for increasing physical activity in youth: A review. *Preventive Medicine* 39(1):157-163.

Kahn, E. B., L. T. Ramsey, R. C. Brownson, G. W. Heath, E. H. Howze, K. E. Powell, E. J. Stone, M. W. Rajab, and P. Corso. 2002. The effectiveness of interventions to increase physical activity: A systematic review. *American Journal of Preventive Medicine* 22(4, Supplement 1):73-107.

Kamath, C. C., K. S. Vickers, A. Ehrlich, L. McGovern, J. Johnson, V. Singhal, R. Paulo, A. Hettinger, P. J. Erwin, and V. M. Montori. 2008. Behavioral interventions to prevent childhood obesity: A systematic review and metaanalyses of randomized trials. *Journal of Clinical Endocrinology and Metabolism* 93(12):4606-4615.

Katz, D. L., M. O'Connell, M. C. Yeh, H. Nawaz, V. Njike, L. M. Anderson, S. Cory, and W. Dietz. 2005. Public health strategies for preventing and controlling overweight and obesity in school and worksite settings: A report on recommendations of the Task Force on Community Preventive Services. *Morbidity and Mortality Weekly Report* 54(RR-10):1-12.

Katz, D. L., M. O'Connell, V. Y. Njike, M. C. Yeh, and H. Nawaz. 2008. Strategies for the prevention and control of obesity in the school setting: Systematic review and meta-analysis. *International Journal of Obesity* 32(12):1780-1789.

Khan, L. K., K. Sobush, D. Keener, K. Goodman, A. Lowry, J. Kakietek, and S. Zaro. 2009. Recommended community strategies and measurements to prevent obesity in the United States. *Morbidity and Mortality Weekly Report* 58(RR07):1-26.

Kropski, J. A., P. H. Keckley, and G. L. Jensen. 2008. School-based obesity prevention programs: An evidence-based review. *Obesity* 16(5):1009-1018.

Lemmens, V. E. P. P., A. Oenema, K. I. Klepp, H. B. Henriksen, and J. Brug. 2008. A systematic review of the evidence regarding efficacy of obesity prevention interventions among adults. *Obesity Reviews* 9(5):446-455.

Matson-Koffman, D. M., J. N. Brownstein, J. A. Neiner, and M. L. Greaney. 2005. A site-specific literature review of policy and environmental interventions that promote physical activity and nutrition for cardiovascular health: What works? *American Journal of Health Promotion* 19(3):167-193.

Mayer, K. 2009. Childhood obesity prevention: Focusing on the community food environment. *Family & Community Health* 32(3):257-270.

Micucci, S., H. Thomas, and J. Vohra. 2002. *The effectiveness of school-based strategies for the primary prevention of obesity and for promoting physical activity and/or nutrition, the major modifiable risk factors for type 2 diabetes: A review of reviews*. Hamilton, ON: Effective Public Health Practice Project.

Moher, D., D. J. Cook, S. Eastwood, I. Olkin, D. Rennie, and D. F. Stroup. 1999. Improving the quality of reports of meta-analyses of randomised controlled trials: The QUOROM statement. *Lancet* 354(9193):1896-1900.

Mulrow, C. D., D. J. Cook, and F. Davidoff. 1997. Systematic reviews: Critical links in the great chain of evidence. *Annals of Internal Medicine* 125(5):389-391.

Peersman, G., S. Oliver, and A. Oakley. 1997. *EPPI-Centre review guidelines*. London: EPPI-Centre, Social Science Research Unit.

Pignone, M. P., A. Ammerman, L. Fernandez, C. T. Orleans, N. Pender, S. Woolf, K. N. Lohr, and S. Sutton. 2003. Counseling to promote a healthy diet in adults: A summary of the evidence for the U.S. Preventive Services Task Force. *American Journal of Preventive Medicine* 24(1):75-92.

Rees, R., J. Kavanagh, A. Harden, J. Shepherd, G. Brunton, S. Oliver, and A. Oakley. 2006. Young people and physical activity: A systematic review matching their views to effective interventions. *Health Education Research* 21(6):806-825.

Reilly, J. J., and Z. C. McDowell. 2003. Physical activity interventions in the prevention and treatment of paediatric obesity: Systematic review and critical appraisal. *Proceedings of the Nutrition Society* 62(3):611-619.

Renzaho, A. M., D. Mellor, K. Boulton, and B. Swinburn. 2010. Effectiveness of prevention programmes for obesity and chronic diseases among immigrants to developed countries: A systematic review. *Public Health Nutrition* 13(3):438-450.

Salmon, J., M. L. Booth, P. Phongsavan, N. Murphy, and A. Timperio. 2007. Promoting physical activity participation among children and adolescents. *Epidemiologic Reviews* 29(1):144-159.

Scottish Intercollegiate Guidelines Network. 2002. A guideline developer's handbook. http://www.sign.ac.uk/guidelines/fulltext/50/index.html (accessed July 1, 2002).

Sharma, M. 2007. International school-based interventions for preventing obesity in children. *Obesity Reviews* 8(2):155-167.

Small, L., D. Anderson, and B. M. Melnyk. 2007. Prevention and early treatment of overweight and obesity in young children: A critical review and appraisal of the evidence. *Pediatric nursing* 33(2).

Stetler, C. B., D. Morsi, S. Rucki, S. Broughton, B. Corrigan, and J. Fitzgerald. 1998. Utilization focused integrative reviews in a nursing service. *Applied Nursing Research* 11:195-206.

Stice, E., H. Shaw, and C. N. Marti. 2006. A meta-analytic review of obesity prevention programs for children and adolescents: The skinny on interventions that work. *Psychological Bulletin* 132(5):667-691.

Summerbell, C. D., E. Waters, L. D. Edmunds, S. Kelly, T. Brown, and K. J. Campbell. 2005. Interventions for preventing obesity in children. *Cochrane Database of Systematic Reviews (Online)* (3).

Thomas, H. 2006. Obesity prevention programs for children and youth: Why are their results so modest? *Health Education Research* 21(6):783-795.

Thomas, H., D. Ciliska, J. Wilson-Abra, S. Micucci, M. Dobbins, and J. Dwyer. 2004. *Effectiveness of physical activity enhancement and obesity prevention programs in children and youth*. Hamilton, Canada: McMaster University and the Public Health and Community Services Department, Public Health Research, Education, and Development Program.

Timperio, A., J. Salmon, and K. Ball. 2004. Evidence-based strategies to promote physical activity among children, adolescents and young adults: Review and update. *Journal of Science and Medicine in Sport/Sports Medicine Australia* 7(1, Supplement):20-29.

Van Sluijs, E. M. F., A. M. McMinn, and S. J. Griffin. 2007. Effectiveness of interventions to promote physical activity in children and adolescents: Systematic review of controlled trials. *British Medical Journal* 335(7622):703-707.

Wareham, N. J., E. M. F. Van Sluijs, and U. Ekelund. 2005. Physical activity and obesity prevention: A review of the current evidence. *Proceedings of the Nutrition Society* 64(2):229-247.

Wofford, L. G. 2008. Systematic review of childhood obesity prevention. *Journal of Pediatric Nursing* 23(1):5-19.

Zenzen, W., and S. Kridli. 2009. Integrative review of school-based childhood obesity prevention programs. *Journal of Pediatric Health Care* 23(4):242-258.

D

Information Sources for Locating Evidence

The following tables provide a listing of selected information sources for locating evidence relevant to obesity prevention.

TABLE D-1 Scientific Literature Databases

Information Source[a]	Description
Academic Search Premier	The world's largest academic multidisciplinary database; coverage spans virtually every area of academic study and includes information dating as far back as 1975.
Agricola	Administered by the National Agricultural Library, indexes more than 2,000 serial titles, as well as books, pamphlets, conference proceedings, research reports, government documents, monographs, theses, patents, software, audiovisual content, and microforms pertaining to all aspects of agriculture, including biotechnology, ecology, entomology, and food and nutrition.
BIOSIS	Coverage includes traditional areas of biology, such as botany, zoology, and microbiology, as well as related fields such as biomedicine, agriculture, pharmacology, and ecology. Interdisciplinary fields such as medicine, biochemistry, biophysics, bioengineering, and biotechnology are also included, as are content summaries, book and meeting abstracts, papers, and posters.
Cochrane Database of Systematic Reviews	Includes the full text of the regularly updated systematic reviews of the effects of health care prepared by the Cochrane Collaboration. Two types of reviews are presented: complete reviews (regularly updated Cochrane Reviews, prepared and maintained by Collaborative Review Groups) and protocols (protocols for reviews currently being prepared [all include an expected date of completion]; protocols are the background, objectives, and methods of reviews in preparation.)
Cochrane Public Health Review Group	Purpose is to produce systematic reviews of population-level interventions, contribute to methodological developments for research synthesis of public health research, transfer knowledge and exchange initiatives, and develop strategies for building primary research and rigorous evaluations of population-level interventions (policy, programs, legislation, community interventions), including maintaining a specialized studies register.

continued

TABLE D-1 Continued

Information Source[a]	Description
Cumulative Index to Nursing and Allied Health Literature (CINAHL)	The most comprehensive resource for nursing and allied health literature; its other indexes include health care books, conference proceedings, standards of professional practice, and audiovisual material.
Exerpta Medica (EMBASE)	Major biomedical and pharmaceutical database indexing more than 3,500 international journals in the following fields: drug research, pharmacology, pharmaceutics, toxicology, clinical and experimental human medicine, health policy and management, public health, occupational health, environmental health, drug dependence and abuse, psychiatry, forensic medicine, and biomedical engineering/instrumentation.
HealthSTAR	Contains citations to the published literature on health services, technology, administration, and research.
MEDLINE	The U.S. National Library of Medicine's (NLM's) premier bibliographic database covering the fields of medicine, nursing, dentistry, veterinary medicine, the health care system, and the preclinical sciences. MEDLINE contains bibliographic citations and author abstracts from more than 4,600 biomedical journals published in the United States and 70 other countries.
National Technical Information Service (NTIS)	The major resource for locating U.S. government–sponsored research reports and studies in the physical sciences, technology, engineering, the biological sciences, medicine and the health sciences, agriculture, and the social sciences.
PsycINFO	Contains citations and summaries of journal articles, book chapters, books, and technical reports, as well as citations to dissertations, all in the field of psychology and psychological aspects of related disciplines.
Science Direct	Encompasses scientific, technical, and medical peer-reviewed journals. Subject coverage includes the biological sciences, business management and accounting, computer science, the earth and planetary sciences, engineering and technology, environmental science, materials science, mathematics, medicine, physics and astronomy, psychology, and social science.
SCOPUS	Coverage includes the life and health sciences; the social sciences; psychology; economics; the biological, agricultural, and environmental sciences; and the general sciences.
Web of Science	Searches 9,200 journals in more than 45 different languages across the sciences, social sciences, and arts and humanities.
WorldCat	Indexes books, conference proceedings, government reports, dissertations and theses, and even online-accessible publications in all scientific, social science, and arts and humanities disciplines.

[a] Note that a subscription may be required to gain full access to some of these databases.

TABLE D-2 Grey and Unpublished Literature

Information Source[a]	Description
Lexis-Nexis	Includes 9,000 newspapers, news magazines, transcripts, the *Congressional Record*, the *Federal Register*, public laws, law review journals, and case law. (A subscription is required.)
New York Academy of Medicine Grey Literature Page	Focuses on grey literature resources from the medical field and includes an extensive listing of agencies and organizations that produce health-related materials. The site also features a quarterly *Grey Literature Report* listing many items that are available free of charge online. (http://www.nyam.org/library/pages/grey_literature_report)
University of Maryland Libraries Virtual Technical Reports Central	Contains links to institutions that produce grey literature, such as technical reports, preprints, reprints, dissertations, theses, and research reports of all kinds. The institutions listed provide either full-text reports or searchable extended abstracts of technical reports on the Internet. (http://www.lib.umd.edu/ENGIN/TechReports/Virtual-TechReports.html)

[a] Note that a subscription may be required to gain full access to some of these databases.

TABLE D-3 Surveys, Polls, and Rankings

Information Source	Description
Behavioral Risk Factor Surveillance System (BRFSS)	The world's largest ongoing telephone health survey system. Operated by the Centers for Disease Control and Prevention (CDC) it has collected data in the United States annually since 1984. The state-based system collects information on health risk behaviors, preventive health practices, and health care access related primarily to chronic disease and injury. States use BRFSS data to identify emerging health problems, establish and track health objectives, and develop and evaluate public health policies and programs. Many states also use BRFSS data to support health-related legislative efforts. (http://www.cdc.gov/brfss/)
Center for Health and Health Care in Schools (CHHCS)	Contains results from parent and teacher public opinion polls and web-based surveys conducted through the CHHCS on a number of topics, including nutrition and physical activity at school. (http://www.healthinschools.org/Publications-and-Resources/Polls-and-Surveys.aspx)
County Health Rankings	Provides access to 50 state reports, ranking each county within the 50 states according to its health outcomes and the multiple factors that determine a county's health. Each county receives a summary rank for its health outcomes and health factors and also for the four different types of health factors: health behaviors, clinical care, social and economic factors, and the physical environment. Each county can also drill down to see specific county-level data (as well as state benchmarks) for the measures upon which the rankings are based. These data could be used to gather information to help answer "Why" questions. (http://www.countyhealthrankings.org/)
Gallup	Gathers poll and survey data regarding a number of health and well-being topics. (http://www.gallup.com)
National Health and Nutrition Examination Surveys (NHANES)	A program of studies designed to assess the health and nutritional status of adults and children in the United States. The survey is unique in that it combines interviews and physical examinations. (http://www.cdc.gov/nchs/nhanes.htm)
U.S. Department of Health and Human Services, Community Health Status Indicators (CHSI)	Provides an overview of key health indicators for local communities to encourage dialogue about actions that can be taken to improve a community's health. The CHSI report was designed not only for public health professionals but also for citizens who are interested in the health of their community. It contains more than 200 measures for each of the 3,141 U.S. counties, and includes data on morbidity and mortality and health behaviors. (http://www.communityhealth.hhs.gov)

TABLE D-4 Government

Information Source[a]	Description
Congressional Research Service (CRS) Reports via National Council for Science in the Environment (NCSE)	NCSE provides CRS reports on a variety of topics concerning the environment and related topics. (http://ncseonline.org/NLE/CRS/)
THOMAS	Provides full-text coverage of the *Congressional Record*, committee reports, and public laws, as well as links to committee home pages and House and Senate directories. (http://thomas.gov)

[a] Note that a subscription is required to gain full access to these databases.

TABLE D-5 Policies and Programs

Information Source	Description	Website
CDC's Guide to Community Preventive Services	A resource designed to aid in the choice of programs and policies to improve health and prevent disease in communities; systematic reviews are used to answer questions about the effectiveness of program/policy interventions, whether an effective intervention is appropriate for a given community, and what effective interventions cost or will return on an investment.	http://www.thecommunityguide.org
CDC's Healthier Worksite Initiative	Provides toolkits containing customizable solutions to help program planners plan, implement, and evaluate a specific health promotion intervention in the categories of general workforce health promotion, healthful eating, physical activity, preventive health screenings, and healthful choices.	http://www.cdc.gov/nccdphp/dnpa/hwi/toolkits/index.htm
CDC's Lean*Works*!	Designed specifically to translate *Community Guide* recommendations into practical tools for employers in the worksite setting. It offers interactive tools and evidence-based resources for designing effective worksite obesity prevention and control programs, including an obesity cost calculator to estimate how much obesity is costing a company and how much a company could save with different workplace interventions.	http://www.cdc.gov/LEANWorks/
Center of Excellence for Training and Research Translation	Focused on providing practitioners with the best available evidence and approaches related to the prevention and control of obesity. Provides resources designed to support the planning, implementation, and evaluation of evidence-supported nutrition, physical activity, and obesity prevention interventions.	http://www.centertrt.org/index.cfm?fa=op.overview
Common Community Measures for Obesity Prevention	Initiated by CDC, identifies a core set of community-level measures focused on policy and environmental strategies that impact obesity prevention. These measures can be used for research and can serve as baseline variables for planning and evaluation.	http://www.cdc.gov/mmwR/PDF/rr/rr5807.pdf.
International City/County Management Association	*Report on Community Health and Food Access: The Local Government Role*	http://www.icma.org/upload/library/2006-09/%7B5CD4101C-2803-4655-9A51-465461B3C897%7D.pdf
Local Government Commission	*Cultivating Community Gardens: The Role of Local Government in Creating Healthy, Livable Neighborhoods*	http://www.lgc.org/freepub/docs/community_design/fact_sheets/community_gardens.pdf
	Healthy Kids, Healthy Communities: School and Local Government Collaborations	http://www.lgc.org/freepub/docs/community_design/fact_sheets/Healthy_Kids_Healthy_Communities.pdf

continued

TABLE D-5 Continued

Information Source	Description	Website
National Academy for State Health Policy	*State Approaches to Childhood Obesity: A Snapshot of Promising Practices and Lessons Learned*	http://www.nashp.org/Files/Obesity_final_with_correct_appendix_C.pdf
National Association of Counties	*County Government Approaches to Combating Youth Obesity, Encouraging Physical Activity, and Creating Healthy Communities* *Transportation Solutions to Create Active, Healthy Counties: Collaboration for Childhood Obesity Prevention*	http://www.naco.org/Content/ContentGroups/Programs_and_Projects/Community_Development/Center_for_Sustainable_Communities/Final_Youth_Obesity_SurveyReport.pdf http://www.naco.org/Template.cfm?Section=New_Technical_Assistance&template=/ContentManagement/ContentDisplay.cfm&ContentID=27724
National Association of County and City Health Officials	*Land Use Planning and Community Design: The Role of Local Public Health Agencies*	http://archive.naccho.org/Documents/Land-use-3.pdf
National Association of Latino Elected Officials	*Obesity in Your Community*	http://www.naleo.org/downloads/HealthToolKit.pdf
National Association of Local Boards of Health	*Land Use Planning for Public Health: The Role of Local Boards of Health in Community Design and Development*	http://www.activeliving.org/files/NALBOH_land_use_report.pdf
National Conference of State Legislatures	*Promoting Healthy Communities and Reducing Childhood Obesity: Legislative Options*	http://www.ncsl.org/default.aspx?tabid=18135
National Governors Association	*Shaping a Healthier Generation: Successful State Strategies to Prevent Childhood Obesity*	http://www.nga.org/Files/pdf/0909HEALTHIERGENERATION.PDF
National League of Cities	Action Kit for Municipal Leaders: Combating Childhood Obesity	http://www.nlc.org/IYEF/PUBLICATIONS___RESOURCES/actionkits.aspx#obesity
Prevention Institute	Environmental Nutrition and Activity Community Tool	http://www.preventioninstitute.org/sa/enact/members/index.php
Robert Wood Johnson Foundation	Health Policy Tracking Service: provides information on important developments in state legislation, policies, and programs affecting health care, primarily through web-based reports and databases. Its primary audience is state legislators and their staffs.	http://www.rwjf.org/pr/product.jsp?id=47788
Robert Wood Johnson Foundation	Healthy Eating Research Results Active Living Research Results	http://www.healthyeatingresearch.org/content/view/12/35/ http://www.activelivingresearch.org/alr/grantsearch/search

TABLE D-5 Continued

Information Source	Description	Website
Trust for America's Health	*F as in Fat: How Obesity Policies Are Failing in America*	http://healthyamericans.org/reports/obesity2008/Obesity2008Report.pdf
U.S. Conference of Mayors	*Mayors' Guide to Fighting Childhood Obesity*	http://usmayors.org/chhs/healthycities/documents/guide-20080326.pdf

E

An In-Depth Look at Study Designs and Methodologies

This appendix provides an in-depth look at study designs and methodologies. It first reviews selected designs (regression discontinuity designs, interrupted time series analysis, observational studies, pre-/posttest designs, and economic cost analysis) commonly used as alternatives to randomized experiments. It concludes with guidance on how to use theory, professional experience, and local wisdom to adapt the evidence gathered to local settings, populations, and times.

COMMON RESEARCH DESIGNS

Reichardt (2006) presents a typology that encompasses the full range of randomized and strong quasi-experimental nonrandomized designs (see Table E-1). This typology is useful because it can substantially broaden the range of design options that can be considered by researchers. Reichardt considers all possible designs that can be created

TABLE E-1 A Typology of Research Designs

Prominent Size-of-Effect Factor	Assignment to Treatment		
		Nonrandom	
	Random	Explicit Quantitative Ordering	No Explicit Quantitative Ordering
Recipients	Randomized recipient design	Regression discontinuity design	Nonequivalent recipients design
Times	Randomized time design	Interrupted time series design	Nonequivalent times design
Settings	Randomized setting design	Discontinuity across settings design	Nonequivalent settings design
Outcome variables	Randomized outcome variable design	Discontinuity across outcome variables design	Nonequivalent outcome variables design

SOURCE: Reichardt, 2006.

based on the combination of two dimensions: assignment rule and primary dimension for assignment of units.

With respect to assignment rule, units can be assigned (1) according to a randomized allocation scheme, (2) on the basis of a quantitative assignment rule, or (3) according to an unknown assignment rule. Randomization schemes in which each unit has an equal probability of being in a given treatment condition are familiar. A quantitative assignment rule means that there is a fixed rule for assigning units to the intervention on the basis of a quantitative measure, typically of need, merit, or risk. For example, organ transplants are allocated on the basis of a weighted combination of patient waiting time and the quality of the match of the available organ to the patient. Finally, unknown assignment rules commonly apply when units self-select into treatments or researchers give different treatments to preexisting groups (e.g., two communities, two school systems). Unknown assignment rules are presumed to be nonrandom.

With respect to units, participants (people or small clusters of people), times, settings, or outcome measures may serve as the units of analysis. Research in public health and medicine commonly assigns treatments to individual (or small groups of) participants. But other units of assignment are possible and should be entertained in some research contexts. Time can be the unit of assignment, as, for example, in some drug research in which short-acting drugs are introduced and withdrawn, or behavior modification interventions are introduced and withdrawn to study their effects on the behavior of single patients. Settings can be the unit of assignment, as when different community health settings are given different treatments, or different intersections are given different treatments (e.g., photo radar monitoring of speeding in a traffic safety study). Finally, even outcome measures can be assigned to different conditions. In a study of the effectiveness of the Sesame Street program, for example, different sets of commonly used letters (e.g., [a, o, p, s] versus [e, i, r, t]) could be selected for inclusion in the program. The knowledge of the specific letters chosen for inclusion in the program could be compared with the knowledge of the control letters to assess the program's effectiveness. Once again, each of these types of units could potentially be assigned to treatment conditions using any of the three assignment rules.

Reichardt provides a useful heuristic framework for expanding thinking about strong alternative research designs. When individuals are not the unit of analysis, however, complications may arise in the statistical analysis. These complications are addressable.

In this section, some commonly utilized quasi-experimental designs from Reichardt's framework are described. First, two designs involving nonrandom, quantitative assignment rules—the regression discontinuity design and the interrupted time series design—are discussed. Next, the observational study (also known as the nonequivalent control group design or nonequivalent recipients design), in which the basis for assignment is unknown, is considered. Finally, the pre-experimental pre–post design, commonly utilized by decision makers, is discussed. For each design, Campbell's and Rubin's perspectives (detailed in Chapter 8) are brought into the discussion as the basis for suggestions for enhancements that may lead to stronger causal inferences.

Regression Discontinuity (RD) Design

Often society prescribes that treatments be given to those with the greatest need, risk, or merit. A quantitative measure is assessed at baseline (or a composite measure is created from a set of baseline measures), and participants scoring above (or below) a threshold score are given the treatment. To cite three examples from the educational arena, access to free lunches is often given to children whose parents have an income below a specified threshold (e.g., the poverty line), whereas children above the poverty line do not receive free lunches. The recognition of dean's list is awarded only to students who achieve a specified grade point average (e.g., 3.5 or greater). And children who reach their sixth birthday by December 31 are enrolled in first grade the following August, whereas younger children are not. Given assessment of the outcome following the intervention, comparison of the outcomes at the threshold for the intervention and in control groups permits strong causal inferences to be drawn.

To understand the RD design, consider the example of evaluating the effectiveness of school lunch programs on health, which is illustrated in Figure E-1. In the figure, all children with a family income of less than $20,000 qualify for the program, whereas children whose families exceed this threshold do not. The outcome measure (here a measure of health problems, such as number of school absences or school nurse visits) is collected for each child. In modeling the relationship between the known quantitative assignment variable (family income) and the outcome, the treatment effect will be represented by the difference in the levels of the regression lines at the cutpoint. In the basic RD design, treatment assignment is determined entirely by the assignment variable. Proper modeling of the relationship between the assignment variable and the outcome permits a strong inference of a treatment effect if there is a discontinuity at the cutpoint.

Ludwig and Miller (2007) used this design to study some of the educational and health effects of the implementation of the original Head Start program in 1965. When the program was launched, counties were invited to submit applications for Head Start funding. In a special program, the 300 poorest counties in the United States (poverty rates exceeding a threshold of 59.2 percent) received technical assistance in writing the Head Start grant proposal. Because of the technical assistance intervention, a very high proportion (80 percent) of the poorest counties received funding, approximately twice the rate of slightly better-off counties (49.2 percent to 59.2 percent poverty rates) that did not receive this assistance. The original Head Start program included not only its well-known educational program, but also basic health services to children (e.g., nutrition supplements and education, immunization, screening). In addition to positive effects on educational achievement, Ludwig and Miller

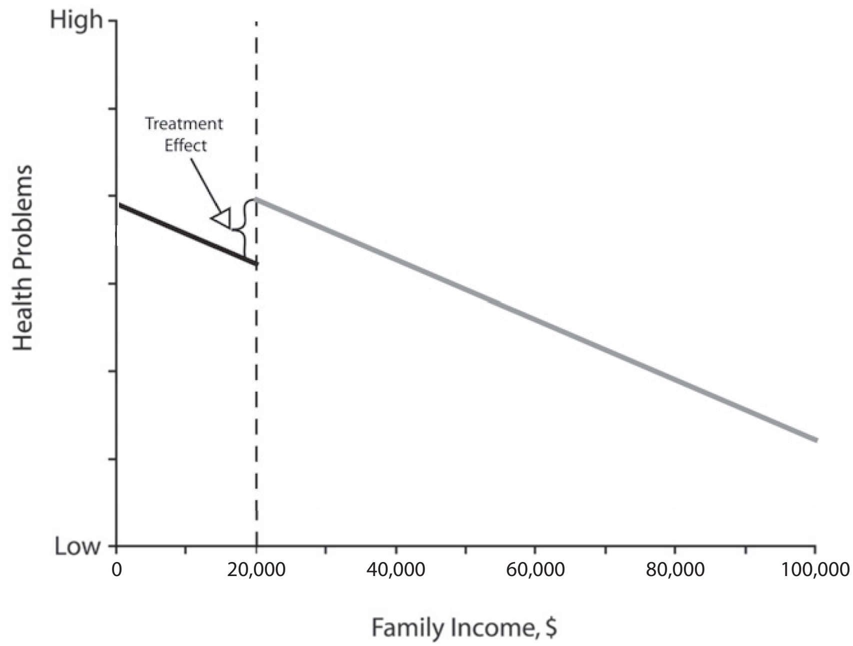


FIGURE E-1 Illustration of the regression discontinuity design using the example of an evaluation of the effect of school lunch programs on children's health.
NOTE: All children whose family income was below the threshold, here $20,000 (dotted line), received the treatment program (school lunch program); all children whose family income was above the threshold did not receive the program. The difference between the regression lines for the program and no-program groups at the threshold represents the treatment effect.
SOURCE: West et al., 2008. Reprinted with permission. West et al., Alternatives to the randomized controlled trial, *American Journal of Public Health*, 98(8):1364, Copyright © 2008 by the American Public Health Association.

found results demonstrating lower mortality rates in children aged 5 to 9 from diseases addressed by the program (e.g., measles, anemia, diabetes).

The RD design overcomes several of the objections to the randomized controlled trial (RCT) discussed in this report. When an existing program uses quantitative assignment rules, the RD design permits strong evaluation of the program without the need to create a pool of participants willing to be randomized. Sometimes outcome data may be collected routinely from large samples of individuals in the program. As illustrated by the Ludwig and Miller study, the design can be used when individual participants, neighborhoods, cities, or counties are the unit of assignment. When new programs are implemented, assignment on the basis of need or risk may be more acceptable to communities that may be resistant to RCTs. The use of a clinically meaningful quantitative assignment variable (e.g., risk level) may help overcome ethical or political objections when a promising potential treatment is being evaluated. Given that the design can often be implemented with the full population of interest—a state, community, school, or hospital—it provides direct evidence of population-level effects.

The RD design is viewed as one of the strongest alternatives to the RCT from both Campbell's (Cook, 2008; Shadish et al., 2002; Trochim, 1984) and Rubin's (Imbens and Lemiuex, 2008; Rubin, 1977) perspectives. However, it introduces two new challenges to causal inference that do not characterize the RCT. First, it is assumed that the functional form of the relationship between the quantitative assignment variable and the outcome is properly modeled. Early work on the RD design in the behavioral sciences typically assumed that a regression equation representing a linear effect of the assignment score on the outcome plus a treatment effect estimated at the cutoff would be sufficient to characterize the relationship. More recent work in econometrics has emphasized the use of alternative methods to characterize the relationship between the assignment variable and the outcome separately above and below the threshold level. For example, with large sample sizes, nonparametric regression models can be fit separately above and below the threshold to minimize any possibility that the functional form of the relationship is not properly specified. Second, in some RD designs, the quantitative assignment variable does not fully determine treatment assignment. Econometricians make a distinction between "sharp" RD designs, in which the quantitative assignment variable fully determines treatment assignment, and "fuzzy" RD designs, in which a more complex treatment selection model determines assignment. These latter designs introduce considerably more complexity, but new statistical modeling techniques based on the potential outcomes perspective (see Hahn et al., 2001) minimize any bias in the estimate of treatment effects.

From Campbell's perspective, several design elements can potentially be used to strengthen the basic design. Replication of the original study using a different threshold can help rule out the possibility that some form of nonlinear growth accounts for the results. Masking (blinding) the threshold score from participants, test scorers, and treatment providers, when possible, can minimize the possibility that factors other than the quantitative assignment variable determine treatment. Investigating the effects of the intervention on a nonequivalent dependent variable that is expected to be affected by many of the same factors as the primary outcome variable, but not the treatment, can strengthen the inference. In the case of fuzzy RD designs, sensitivity analyses in which different plausible assumptions are made about alternative functional forms of the relationship and selection models can also be conducted.

Interrupted Time Series (ITS) Analysis

Often policy changes go into effect on a specific date. To illustrate, the Federal Communications Commission (FCC) allowed television broadcasting to be introduced for the first time in several medium-sized cities in the United States in 1951. Bans on indoor smoking have been introduced in numerous cities (and states) on specific dates. If outcome data can be collected or archival data are available at regular fixed intervals (e.g., weekly, monthly), the ITS provides a strong design for causal inference. The logic of the ITS closely parallels that of the RD design except that the threshold on the time rather than the baseline covariate is the basis for treatment assignment (Reichardt, 2006).

Khuder and colleagues (2007) present a nice illustration of an ITS. An Ohio city instituted a ban on smoking in indoor workplaces and public places in March 2002. All cases of angina, heart failure, atherosclerosis, and acute myocardial infarction in city hospitals were identified from hospital discharge data. Following the introduction of the smoking ban, a significant reduction of heart disease–related hospital admissions was seen.

From a design standpoint, causal inferences from the simple ITS perspective need to be tempered because the basic design fails to address three major threats to the certainty of the causal relationship between an intervention and the observed outcomes (internal validity) (Shadish et al., 2002; West et al., 2000). First, some other confounding event (e.g., introduction of a new heart medication) may occur at about the same time as the introduction of the intervention. Second, some interventions may change the population of participants in the area. For example, some cities have offered college scholarships to all students who graduate from high school. In such cases, in addition to any effect of the program on the achievement of city residents, the introduction of the program may foster immigration of highly education-oriented families to the city, changing the nature of the student population. Third, record-keeping practices may change at about the time of the intervention. For example, new criteria for the diagnosis of angina or myocardial infarction may change the number of heart disease cases even in the absence of any effect of the intervention.

From a statistical standpoint, several potential problems with longitudinal data need to be addressed. Any long-term natural trends (e.g., a general decrease in heart disease cases) or cycles (e.g., more admissions during certain seasons of the year) in the data need to be modeled so their effects can be removed. In addition, time series data typically reflect serial dependence: observations closer in time tend to be more similar than observations further apart in time. These problems need to be statistically modeled to remove their effects, permitting proper estimates of the causal effect of the intervention and its standard error. Time series analysis strategies have been developed to permit researchers to address these issues (e.g., Chatfield, 2004). In addition, as in the RD design, the actual introduction of the intervention may be fuzzy. In the Khuder et al. study, for example, there was evidence that the enforcement and full implementation of the smoking ban required some months after the ban was enacted. In such cases, a function describing the pattern of implementation of the intervention may need to be included in the model (e.g., Hennigan et al., 1982).

In the Campbell tradition, causal inferences drawn from the basic ITS design can be greatly strengthened by the addition of design elements that address threats to validity. Khuder and colleagues included another, similar Ohio city that did not institute a smoking ban (control series), finding no parallel change in heart disease admissions after the March 2002 timepoint when the smoking ban was introduced in the treatment city. They also found that hospital admissions for diagnoses unrelated to smoking did not change in either city after March 2002 (nonequivalent dependent variable). In some time series applications, a design element known as switching replications can be used, which involves locating another, similar city in which the intervention was introduced at a different timepoint. In their study of the introduction of television and its effects on crime rates, for example, Hennigan and colleagues (1982) located 34 medium-sized cities in which television was introduced in 1951 and 34 cities matched for region and size in which television was introduced in 1954 following the lifting of a freeze on new broadcasting licenses by the FCC. They found a similar effect of the introduction of television on crime rates (e.g., an increase in larceny) beginning in 1951 in the prefreeze cities and in 1954 in the postfreeze cities. In both the Khuder et al. and Hennigan et al. studies, the addition of the design element greatly reduced the likelihood that any threat to the level of certainty of the causal inference (internal validity) could account for the results obtained. As with the RD design, moreover, the ITS design can often be implemented with the full population of interest so that it provides direct evidence of population-level effects.

Observational Studies

The observational study (also known as the nonequivalent control group design) is a quasi-experimental design that is commonly used in applied research on interventions, likely because of its ease of implementation. In this design, a baseline measure and a final outcome measure are collected on all participants. Following the baseline measurement, one group is given the treatment, while the second, comparison group does not receive the treatment. The groups may be preexisting (e.g., schools, communities), or unrelated participants may self-select into the treatment in some manner. For example, Roos and colleagues (1978) used this design to compare the health outcomes of children who received and did not receive tonsillectomies in a province of Canada. The bases on which the selection into the tonsillectomy treatment occurred were unknown and presumed to be nonrandom, possibly depending on such factors as the child's medical history, the family, the physician, and the region. The challenge of this design is that several threats associated with possible interactions between selection and other threats to level of certainty (internal validity) might be plausible. These threats must be addressed if strong causal inferences are to be drawn.

To illustrate this design, consider an evaluation of a campaign to increase sales of lottery tickets (Reynolds and West, 1987). State lottery tickets are sold primarily in convenience stores and contribute to general state revenue or revenue for targeted programs (e.g., education) in several states. The stores refused randomization, ruling out an RCT. The Arizona lottery commission wished to evaluate the effectiveness of a sales campaign to increase lottery ticket sales in an 8-week-long lottery game. In the "Ask for the Sale" campaign, store clerks were instructed to ask each adult customer during checkout if he or she wished to purchase a lottery ticket. A nearby sign informed customers that if the sales clerk did not ask them, they would get a lottery ticket for free.

From Campbell's perspective, four threats to the level of certainty (internal validity) in this example observational study could interact with selection and undermine causal inference regarding the intervention and the observed outcomes (Shadish et al., 2002):

- Selection × history interaction—Some other event unrelated to the treatment could occur during the lottery game that would affect sales. Control stores could be disproportionately affected by nearby highway construction, for example, resulting in a decrease in customer traffic.
- Selection × maturation interaction—Sales in treatment stores could be growing at a faster rate than sales in control stores even in the absence of the treatment.
- Selection × instrumentation interaction—The nature of the measurement of lottery ticket sales in the treatment stores could change during the game in the treatment but not the control group. For example, stores in the treatment group could switch disproportionately from manual reporting of sales to more complete computer recording of sales.
- Selection × statistical regression interaction—Stores having unusually low sales in the previous lottery game could self-select to be in the treatment group. Sales could return to normal levels even in the absence of the intervention.

Reynolds and West (1987) matched treatment and control stores on sales in the prior game and on ZIP code (a proxy for neighborhood socioeconomic status). As shown in Figure E-2, they implemented the basic observational study design and then added several design features to address possible threats to the certainty of the causal relationship (internal validity). Panel (a) displays the results of the basic design, showing no difference in sales in the prior game (game 10, no program intervention in both stores, i.e., baseline) but greater sales in the treatment stores during the campaign (game 11) compared with matched control stores with no program intervention. Panel (b) displays the results from a set of nonequivalent dependent measures, sales categories that would be expected to be affected by other general factors that affect sales but not by the intervention. The increase in sales of lottery tickets was greater than the increase for other sales categories. Panel (c) displays the results from a short time series of observations in which the sales campaign was implemented in the treatment stores in week 4 of the game. The results show that both the treatment and control stores experienced similar levels of sales each week prior to the intervention, but that the treatment stores sold far more tickets each week following the initiation of the campaign at the beginning of week 5 (i.e., "program started" in Figure E-2). Taken together, the pattern of results presented by the basic design and the additional design elements provided strong support for the effectiveness of the sales campaign.

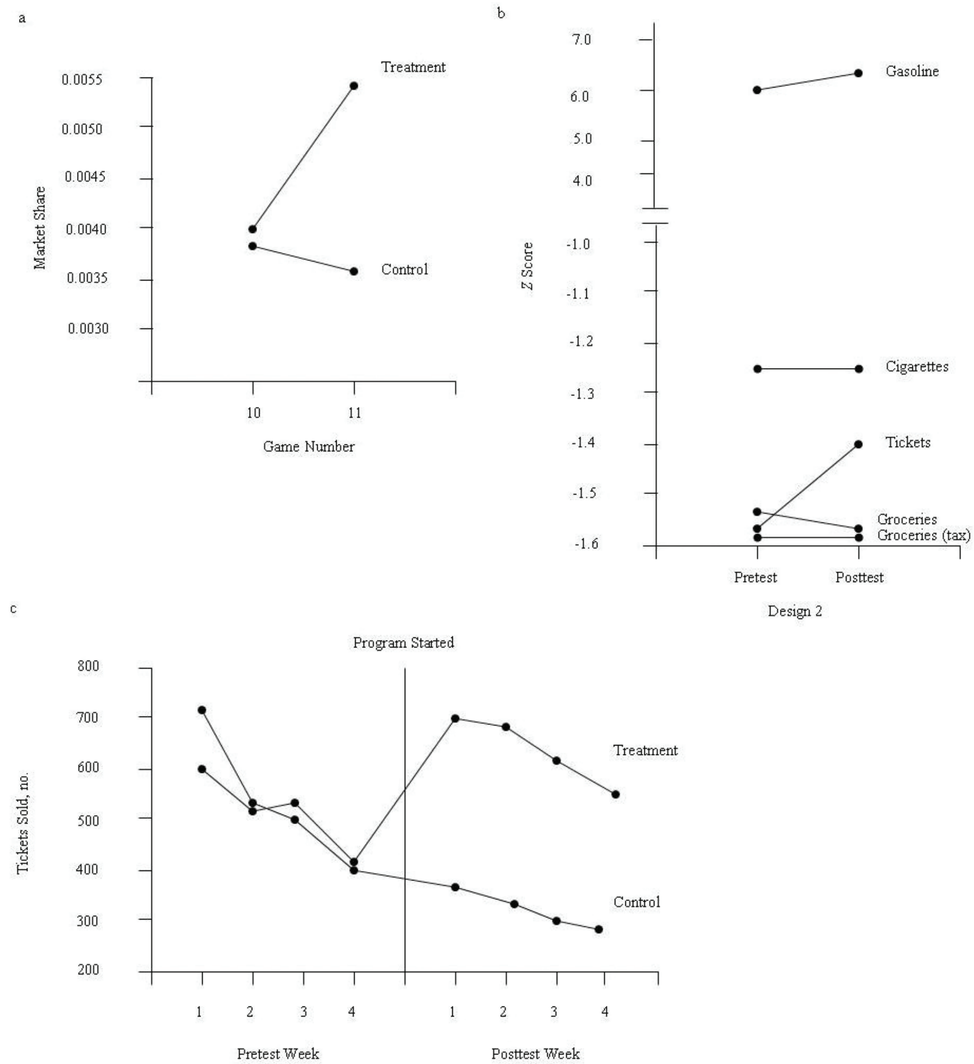


FIGURE E-2 Design elements that strengthen causal inferences in observational studies.
NOTE: (a) Matching. Treatment and control stores were selected from the same chain, were in the same geographic location, and were comparable in sales during the baseline (lottery game 10). Introduction of the treatment at the beginning of lottery game 11 yielded an increase in sales only in the treatment stores. (b) Nonequivalent dependent variables. Within the treatment stores, sales of lottery tickets increased substantially following the introduction of the treatment. Sales of other major categories (gasoline, cigarettes, nontaxable groceries, and taxable groceries) that would be expected to be affected by confounding factors but not the treatment did not show appreciable change. (c) Repeated pre- and posttest measurements. Treatment and control stores' sales showed comparable trends during the 4 weeks prior to the introduction of the treatment. The level of sales in the treatment and control stores was similar prior to the introduction of the treatment, but differed substantially beginning immediately after the treatment was introduced.
SOURCE: Adapted from Reynolds and West, 1987. Reynolds, K. D., and S. G. West, *Evaluation Review* 11:691-714, Copyright © 1987 by SAGE Publications. Reprinted by permission of SAGE Publications.

Rubin's (2006) perspective takes an alternative approach, attempting to create high-quality matches between treatment and control cases on all variables measured at baseline. Two recent developments in the technology of matching procedures have greatly improved the quality of matches that can be achieved. Historically, the number of variables on which high-quality matches could be achieved was limited because of the difficulty of finding cases that matched on several different variables. Rosenbaum and Rubin (1983) developed the idea of matching on the propensity score—the probability that the case would be in the treatment group based on measured background variables. They proved mathematically that if close matches between cases in the treatment and control groups could be achieved on the propensity score, the two groups would also be closely balanced on all baseline variables from which the propensity score was constructed. Propensity scores are typically created on the basis of a logistic regression equation in which the probability of being in the treatment group based on the measured covariates is estimated (see Schafer and Kang, 2008, for a discussion of available techniques). The distributions of each separate baseline covariate in the treatment and control groups can be compared to ensure that close matches have been achieved. Second, new optimal matching algorithms for pairing cases in the two groups maximize the comparability of the groups' propensity scores (Rosenbaum, 2002). The central challenge for these modern matching procedures is ensuring that all important covariates related to both the treatment group and the outcome have been included in the construction of the propensity score. Otherwise stated, have "identical" units been created by the matching procedures?

Wu and colleagues (2008a,b; see also West and Thoemmes, 2010) used these procedures in a study of the effect of retention in first grade on children's subsequent math and reading achievement. Based on input from subject matter experts, they measured 72 variables at baseline believed to possibly be related to retention, achievement, or both. Propensity scores were constructed using logistic regression. From 784 children who were below the median at school entrance, 97 pairs that were closely matched on the propensity scores were constructed using optimal matching procedures. Figure E-3 shows that the distribution of the retained and promoted groups on propensity scores became closely balanced following the use of the optimal matching procedures. The use of propensity scores helps rule out the possibility that preexisting differences (selection bias) between the groups could account for the results obtained.

The central challenge of the propensity score procedure is ensuring that the groups have been closely balanced on all important covariates. For example, matching on a few demographic variables is unlikely to represent fully all of the important baseline differences between the groups (Rubin, 2006). Checks on the balance for each individual covariate can be performed, and propensity scores can be reestimated or specific controls for unbalanced covariates included in the statistical models. Finally, researchers can conduct sensitivity analyses to explore how much the results would change if important covariates were omitted from the propensity score model.

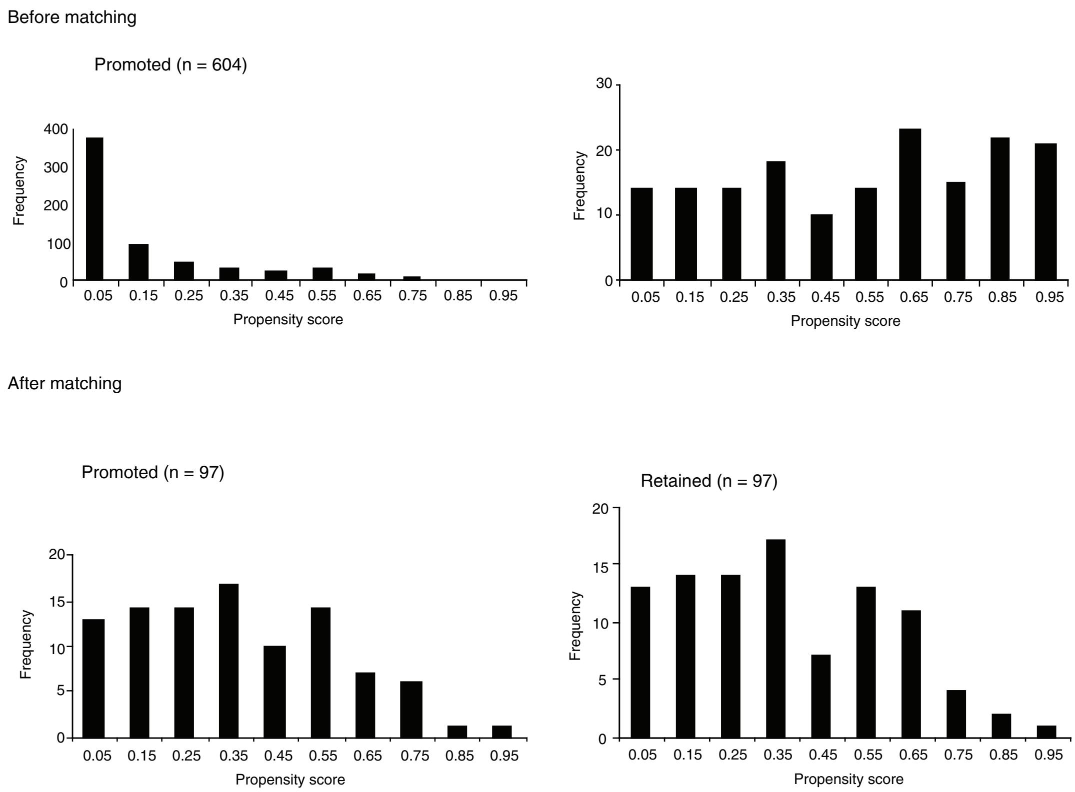


FIGURE E-3 The distributions of propensity score before and after matching in a study of the effects of retention in first grade on children's subsequent math and reading achievement.
NOTE: The scale of the y-axis (frequency) differs between the two top panels (before matching), but is identical for the two bottom panels (after matching). SD = standard deviation.
SOURCE: Wu et al., 2008b.

Pre-/Posttest Designs

Sometimes researchers believe that designs involving a pretest (baseline), a treatment, and a posttest (outcome) measure are sufficient to evaluate the effects of interventions. Such designs are attractive because of their ease of implementation, but they are far weaker in terms of causal inference than experimental and quasi-experimental designs. Campbell and Stanley (1966) present these designs in detail, considering them in a separate, weaker class of "pre-experimental" designs. The following threats to the level of certainty (internal validity) characterize the pretest–posttest design:

- *History*—Some other event unrelated to the treatment occurs between the pretest and posttest that could cause the observed effect in the absence of a treatment effect.
- *Maturation*—Maturational growth or decline in individuals (or secular trends in larger units) could cause the observed effect in the absence of a treatment effect.

- *Instrumentation*—A change in the measuring instrument could lead to increases or decreases in the level of the outcome variable in the absence of a treatment effect. Sometimes these changes can be quite subtle, as when a well-validated measurement instrument may measure different constructs in children of different ages (e.g., standardized math tests).
- *Testing*—Simply the taking of a test or knowledge of the specific dimensions of one's behavior that are being monitored could lead to increases or decreases in scores in the absence of a treatment effect.
- *Statistical regression*—Participants could be selected for treatment that is extreme relative to the mean of their group or is initiated at particular crisis periods in their lives (e.g., a period of binge eating during the holidays). Unreliability in measurement or the passage of the crisis with time could lead to scores that are closer to the mean upon retesting in the absence of a treatment effect. Again, these effects can be quite subtle (see Campbell and Kenny, 1999, for a review).

Any pretest–posttest design may be subject to plausible versions of one or more of these threats to its level of certainty (internal validity) depending on the specific research context, making it very difficult to conclude that a causal effect of the treatment has occurred. Once again, following Campbell's tradition, the strategy of adding design elements to address specific plausible threats can help provide more confidence that the treatment rather than other confounding factors has had the desired effect. For example, replicating the pretest–posttest design at different times with different cohorts of individuals can help rule out history effects, while taking several pretest measures to estimate the maturation trend in the absence of treatment can help rule out history. Box 8-4 in Chapter 8 identifies many of the design elements that can be employed; Shadish and Cook (1999) and Shadish and colleagues (2002) present fuller discussions. The design element approach does not enable the certainty about causal inference provided by the RCT, but it can often greatly improve the evidence base on which decision makers make choices about implementing interventions.

Economic Cost Analysis

Studies that assess the economic costs of obesity can differ in terms of their breadth and perspective. Differences in breadth will be reflected in choices of the population(s) covered (e.g., defined by age, gender, race/ethnicity, socioeconomic status), the range of diseases considered, and the types of costs to include. These decisions will be driven by the perspective taken in the study, as well as by available data.

One key distinction is in the methodological approach employed. In conducting economic cost studies, researchers choose between a "prevalence-based" approach and an "incidence-based" approach (Lightwood et al., 2000). The prevalence based approach (also referred to as an "annual cost" or "cross-sectional" approach) is used to estimate the costs associated with a given problem at a point in time (e.g., a given year), reflecting not just the costs resulting from the condition of interest at that point in time, but also the costs resulting from cumulative past health status variables that manifest themselves at that point in time. Given the lags between poor diets, physical inactivity, obesity, and resulting diseases, cost estimates based on the prevalence approach reflect historical behavior patterns and cannot be used to predict the short-term impact of interventions aimed at reducing obesity and its consequences. Nevertheless, this is the most common approach used in economic cost studies given its relatively simpler methodology and the availability of necessary data.

The incidence-based approach (also referred to as a "lifetime cost" or "longitudinal" approach) aims to estimate the additional costs expected to result from a given condition in a specific population over their lifetimes. When applied to health care costs resulting from obesity, this approach balances the additional health care costs an obese individual faces at a point in time against the health care cost savings that accrue as a result of the shorter lifetime of an obese individual. These estimates can be useful for decision makers, who can estimate the change in costs of obesity that results from policy and other interventions aimed at changing the behaviors that result in obesity. However, the estimates produced by the incidence-based approach can be quite sensitive to assumptions about future costs, changes in health care delivery and technology, and the way future costs are discounted. An additional challenge relates to adequately controlling for the variety of other determinants of costs, that is, trying to estimate costs for a nonobese person so that the true excess costs resulting from obesity can be determined.

Finally, decisions about what costs to include under either approach will have a significant impact on the resulting estimates (Lightwood et al., 2000). There are several key distinctions here, with choices related to the perspective from which one is assessing costs. Economists, for example, distinguish between internal and external costs, with this distinction reflecting who ultimately bears the costs. Internal costs are those that are borne by the individual who engages in a given behavior or has a given condition; for obese individuals, these include the out-of-pocket costs and health insurance premiums they pay for health care, as well as the lower wages they receive if these wages reflect their lower productivity. External costs are those that are borne by individuals as a result of others' behaviors; for example, monies spent on public health insurance programs that cover the costs of treating diseases caused by obesity are shared by all taxpayers, including those at healthy weights. Many argue that only external costs should be considered when making decisions about whether policy interventions are warranted. However, to the extent that other market failures exist (e.g., information failures reflecting a poor understanding of the link between sugar-sweetened beverage consumption and weight that can be further distorted by effective marketing of such beverages or individuals' time-inconsistent preferences that result in short-term gratification at the cost of long-term harm), it may be appropriate to include some portion of internal costs.

A second key distinction relates to a focus on the "gross" versus the "net" costs of a given behavior. The prevalence-based approach produces an estimate of gross costs—those that result from the consequences of obesity in a given year, for example. In contrast, the incidence-based approach produces an estimate of net costs, reflecting the trade-offs between higher average annual costs for an obese individual and the extra costs that result from a nonobese individual's living longer. When one is comparing gross and net costs, a particularly controversial issue relates to what has come to be known as the "death benefit," that is, whether the "savings" that result from lower pension and social security payments that an obese individual who dies prematurely will not collect should be included in the cost accounting.

MATCHING, MAPPING, POOLING, AND PATCHING

Despite best efforts to amass available evidence, decision makers grappling with an emergent problem such as obesity will face inevitable decisions that must be made and actions that must be taken in the relative absence of evidence. This section offers useful methods for combining evidence with theory, expert opinion, experience, and local wisdom about local traditions and probable responses to proposed actions in a process that has been described as "matching, mapping, pooling, and patching" (Green and Glasgow, 2006; Green and Kreuter, 2005, pp. 197-203).

Matching

Matching refers to aligning the source of evidence with the targets of an intervention. The evidence from different disciplines of science and research is distributed according to the level (individual, family, organization, or community) at which it was generated and matched to the level(s) of the proposed intervention and its intended impact (McLeroy et al., 1988; Richard et al., 1996; Sallis et al., 2008). Evidence from psychology and medicine generally focuses on the individual, while evidence from sociology and public health generally applies to groups, organizations, and populations. The ecological thinking and recognition of social determinants of health implicit in this approach have seen a renaissance in public health in recent decades (Berkman and Kawachi, 2000; Best et al., 2003; French et al., 2001; Richard et al., 1996; Sallis et al., 2008). The usual representation of the ecological model in public health and health promotion is a set of concentric circles or ovals, as illustrated in Chapter 1 (see Figure 1-5) (e.g., Booth et al., 2001; Green and Kreuter, 2005, p. 130; IOM, 2007; National Committee on Vital and Health Statistics, 2002). For example, an ecological approach to planning a comprehensive childhood obesity control intervention would involve matching evidence of an intervention's effectiveness with each ecological level to which the evidence might apply:

- evidence from studies of individual change in diet and physical activity, particularly from clinical studies and evaluations of obesity control programs directed at individuals and small groups;
- evidence from studies of family, community, and cultural groups regarding group or population changes, and evaluations of change in response to mass communication or organizational or environmental changes; and
- evidence from policy studies, public health evaluations, and epidemiological studies—for example, the impact of school lunch programs on food consumption patterns, or the impact of policies, executive decisions, and regulatory enforcement of codes for construction of sidewalks or hiking or bike paths on community physical activity patterns.

The preponderance of evidence, especially that judged worthy of inclusion in systematic reviews for evidence-based practice guidelines, tends to be derived from studies of the impact of interventions on individuals. Yet there is widespread agreement that further progress in controlling obesity will require policy changes, organizational changes, changes in the built environment, and changes in social norms, all of which require interventions and measurement of change at levels beyond the individual.

An example of an ecological model that illustrates the matching process is the Multilevel Approach to Community Health (MATCH) model of B. G. Simons-Morton and colleagues (1995) (see Figure E-4). The MATCH model suggests the alignment of evidence with each of the four levels shown by the vertical arrangement of boxes in the figure. It grew out of a conceptualization of intervention research (Parcel, 1987; D. G. Simons-Morton et al., 1988a), and was applied in a series of Intervention Handbooks published by the Centers for Disease Control and Prevention (CDC), including one on promoting physical activity among adults (D. G. Simons-Morton et al., 1988b), one on a series of reviews on health-related physical fitness in childhood (B. G. Simons-Morton et al., 1988a), and one on implementing organizational changes to promote healthful diet and physical activity in schools (B. G. Simons-Morton et al., 1988b). Other ecological models have been suggested for planning and evaluating overall programs (e.g., Green and Kreuter, 2005; Richard et al., 1996) or components of programs such as mass communications (Abroms and Maibach, 2008), and for integrating knowledge for community partnering (Best et al., 2003). The MATCH model is a first step in blending evidence from various sources for the design of a comprehensive intervention that will address the various levels at which evidence is needed and for which adaptations and innovations will be necessary when the evidence is lacking at specific levels.

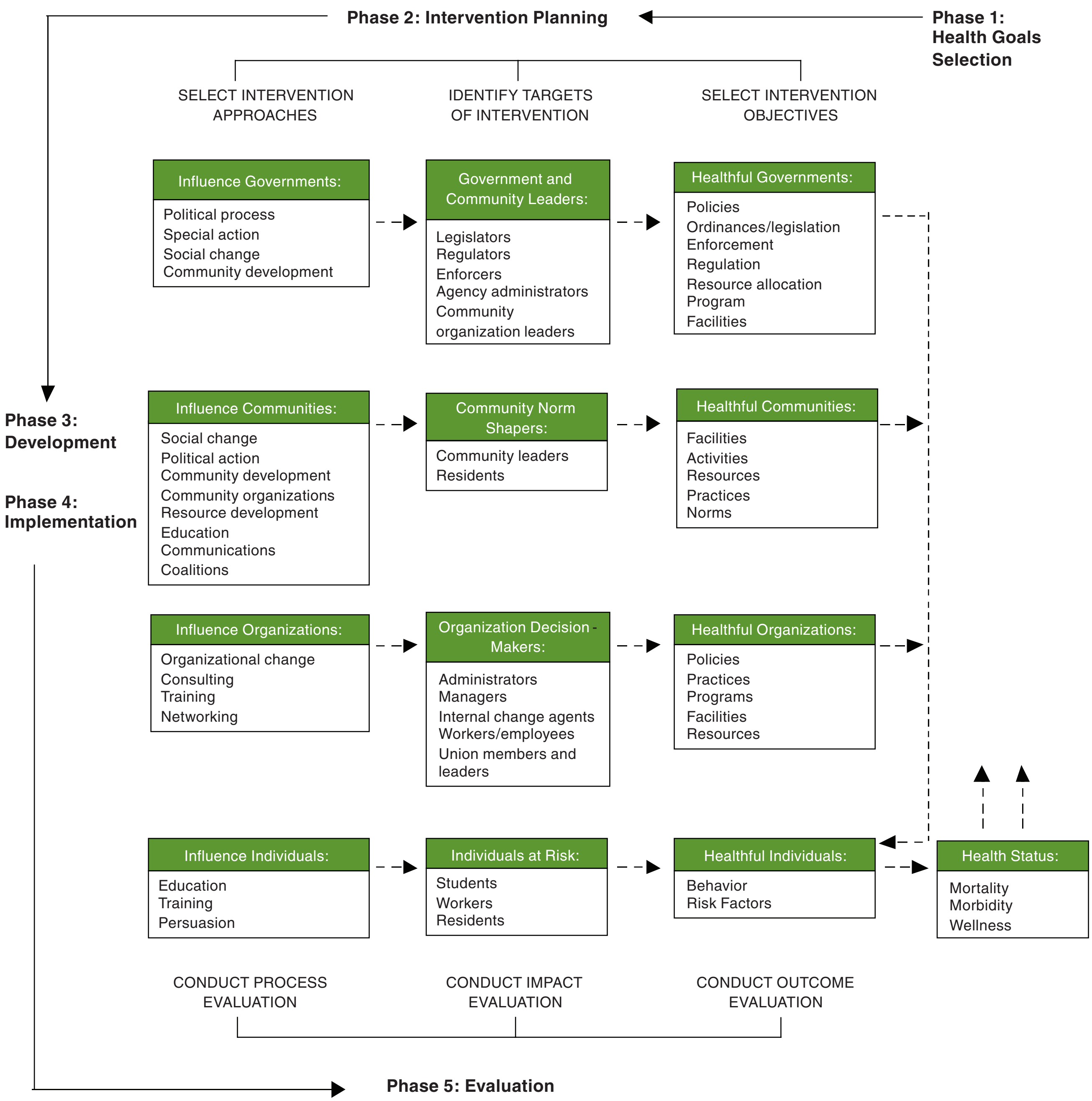


FIGURE E-4 The Multilevel Approach to Community Health (MATCH) model used to align the source of evidence with the targets of an intervention, plan and evaluate programs, and integrate knowledge for community partnering.
SOURCE: B. G. Simons-Morton et al., 1995. Reprinted by permission of Waveland Press, Inc. from B. G. Simons-Morton, W. H. Greene, and N. H. Gottlieb, *Introduction to health education and promotion*, 2nd ed. Long Grove, IL: Waveland Press, Inc., 1995, All rights reserved.

Mapping

Mapping refers to tracing the causal chains or mechanisms of change inferred when evidence is matched with levels of intervention and change. The evidence will be incomplete at each ecological level with respect to the local or state circumstances in which a decision and action must be taken, but theory can (and will, formally or informally, consciously or unconsciously) be brought to bear. People naturally impose their own assumptions about causes and needed actions, and these constitute, in the crudest sense, theories of the problem and the solution. A formal process of mapping objectives, evidence, and formal theories of behavior and social change to the design of coherent, practical interventions includes guidelines for how to assess evidence and theories (Bartholomew et al., 2006). For example, Schaalma and Kok (2006) describe their development of an HIV prevention program using theory to guide their selection of determinants of behavior and the environment (e.g., knowledge, risk perceptions, attitudes, social influences, self-efficacy) relevant to the health problem (e.g., safe sex). Using this process of mapping the objectives of the intervention, existing evidence (both qualitative and quantitative), and theory, the authors were able to identify theoretical methods of behavior change (e.g., social cognitive theory) and translate them into practical strategies (e.g., video-guided role playing). The mapping of theory helps the decision maker fill gaps in the evidence and consider how the evidence from a distant source and a particular population may or may not apply to the local setting, circumstances, and population.

Too often, the published literature in a discipline will provide little if any detail about the theory on which an intervention was based, sometimes because the study was not designed to test a theory, sometimes out of reluctance or inability to be explicit about the theory, and sometimes because editors of health science journals prefer to devote space to methodological detail and data analysis rather than theory. Greenhalgh and colleagues (2007, p. 861) suggest shifting "the balance in what we define as quality from an exclusive focus on empirical method (the extent to which authors have adhered to the accepted rules of controlled trials) to one that embraces theory (the extent to which a theoretical mechanism was explicitly defined and tested)."

Although theory may not be the interest of the end users of evidence, they will, as noted, have their own assumptions in the form of tacit theories. This step will benefit from their consultation with those who have experience with the problem and understand the scientific literature and its formally tested theories of causation and change.

Pooling

Much of the published evidence in obesity-related research is epidemiological or observational, linking the intermediate causes or risk factors with obesity-related health outcomes. Therefore, the range of values of moderating variables (demographic, socioeconomic, and other social factors that modify the observed relationships) in models that predict outcomes is limited. The experimental evidence is even more scarce and limited in generalizability. Because of these limitations of the evidence-based practice literature, decision makers must turn to practice-based evidence and ways of pooling evidence from various extant or emerging programs and practices.

Pooling refers to consultation with decision makers who have dealt with the problem of obesity in a similar population or setting. After matching and mapping the evidence and theories, local decision makers will still be uncertain about how well the evidence applies to each of the mediators (i.e., mechanisms or intermediate steps) and moderators (i.e., conditions that make an association stronger or weaker) in their logic model for local action. At this point, they should turn to the opinions of experts and experienced practitioners in their or similar settings (e.g., Banwell et al., 2005; D'Onofrio, 2001). Methods exist for pooling these opinions and analyzing them in various systematic and formal or unsystematic and informal ways. For example, Banwell and colleagues (2005) used an adapted Delphi technique (the Delphi Method, described in Chapter 6) to obtain views of obesity, dietary, and physical activity experts about social trends that have contributed to an obesogenic environment in Australia. Through this semistructured process, they were able to identify trends in expert opinion, as well as rank the trends to help inform public policy.

Practice-based evidence is that which comes primarily from practice settings, in real time, and from typical practitioners, as distinct from evidence from more academically controlled settings, with highly trained and supervised practitioners conducting interventions under strict protocols. Such tacit evidence, often unpublished, draws on the experience of those who have grappled with the problem and/or the intervention in situations more typical of those in which the evidence would be applied elsewhere. Even when evidence from experimental studies is available, decision makers often ask, understandably, whether it applies to their context—in their practice or policy setting, circumstances, and population (Bowen and Zwi, 2005; Dobbins et al., 2007; Dobrow et al., 2004, 2006; Green, 2008). They want to weigh what the experimental evidence shows, with its strong level of certainty of the causal relationship between the intervention and the observed outcomes (internal validity), against what the experience of their own and similar practices and practitioners has been, with its possibly stronger generalizability (external validity).

Finally, the use of pooling in weighing and supplementing evidence becomes an important negotiating process among organizations cooperating in community-level and other broad collaborative programs and policies. Each participant in such collaborations will weigh different types of evidence differently, and each will have an idiosyncratic view of its own experience and what it says about the problem and the proposed solutions (Best et al., 2003). This recognition of complexities in the evidence and multiplicities of experience has led to a growing interest in systems theory or systems thinking (Green, 2006) (see Chapter 4).

Patching

Engaging people from the community affected by the problem of obesity or its determinants allows for the inclusion of local wisdom from the outset in the adaptation of evidence-based practices and the creation of new ideas to be tested. As with any federal or state health program that must depend on local and state initiative and implementation, the process of rolling out a policy or taking a program to scale requires the engagement and participation of practitioners and populations at the front lines (Ottoson et al., 2009). Success will depend on continuous adaptation to their perception of needs; their understanding of and access to local resources; their willingness, skill, and confidence to implement the recommended intervention; and the reinforcement they will get from doing so (Ottoson and Green, 1987). Many of the methods for patching together evidence-based practices, theory-based programs, and practice-based experiences into a viable effort at the state or local level are contained in manuals and guidelines such as CDC's Planned Approach to Community Health (PATCH; Kreuter, 1992) and Racial and Ethnic Approaches to Community Health, (REACH; CDC, 2009b); community-based participatory research (CBPR; Cargo and Mercer, 2008; Horowitz et al., 2009; Minkler and Wallerstein, 2008); the National Cancer Institute (NCI) and Substance Abuse and Mental Health Services Administration's (SAMHSA) Cancer Control PLANET (NCI and SAMHSA, 2009); and other web resources that need to be made more interactive and responsive as the evidence changes (e.g., http://www.cdc.gov/nutrition/professionals/researchtopractice/index.html [CDC, 2009a]).

An obesity-related example of engaging community members in the planning, delivery, and evaluation of interventions is the Shape Up Somerville environmental change intervention, designed to prevent obesity in culturally diverse, high-risk early elementary school children in Somerville, Massachusetts. One outcome of this initiative was a decrease in BMI *z*-scores among children at high risk for obesity in grades 1-3, a result of an intervention that aimed to bring participants' energy equation into balance by modifying the school, home, and community environments to increase both physical activity options and the availability of healthful foods (Economos et al., 2007). Community members and groups engaged in the intervention included children, parents, teachers, school food service providers, city departments, policy makers, health care providers, restaurants, before- and after-school programs, and the media. Another strategy of Shape Up Somerville was to improve school food service, which led to changes that enhanced the nutrient profiles of and attitudes toward school meals. The engagement of students, parents, teachers, school leaders, and food service personnel was an integral part of the process (Goldberg et al., 2009). There is also a growing literature on how to combine systematic reviews of quantitative and qualitative evidence with realist reviews of theoretical assumptions and with the practical experience of those who must make the final decisions on local action (Ogilvie et al., 2005). For example, Mays and colleagues (2005, p. 7) offer:

- "a description of the main stages in a systematic review of evidence from research and nonresearch sources designed to inform decision making by policy makers and managers;
- an indication of the range of evidence that could potentially be incorporated into such reviews;
- pragmatic guidance on the main methodological issues . . . , given the early stage of development of methods of such reviews, and with a particular focus on the synthesis of qualitative and quantitative evidence;
- an introduction to some of the approaches available to synthesize these different forms of evidence; and
- an indication of the types of review questions particular approaches to synthesis are best able to address."

These alternatives to evidence-based practice of the most literal and rigorous scientific variety suggest some advantages and complementarities of a model of practice-based evidence that produces locally adapted and prospectively tested evidence. Users of this and other guidance for linking research to the decisions they must make in their own settings will need to trade off some degree of rigor for more reality in the setting, conditions of practice, and free-living populations observed.

REFERENCES

Abroms, L. C., and E. W. Maibach. 2008. The effectiveness of mass communication to change public behavior. *Annual Review of Public Health* 29:219-234.

Banwell, C., S. Hinde, J. Dixon, and B. Sibthorpe. 2005. Reflections on expert consensus: A case study of the social trends contributing to obesity. *European Journal of Public Health* 15(6):564-568.

Bartholomew, L. K., G. S. Parcel, G. Kok, and N. H. Gotleib. 2006. *Planning health promotion programs: An intervention mapping approach*. San Francisco, CA: Jossey-Bass Wiley.

Berkman, L. F., and I. Kawachi, eds. 2000. *Social epidemiology.* New York: Oxford University Press.

Best, A., D. Stokols, L. W. Green, S. Leischow, B. Holmes, and K. Buchholz. 2003. An integrative framework for community partnering to translate theory into effective health promotion strategy. *American Journal of Health Promotion* 18(2):168-176.

Booth, S. L., J. F. Sallis, C. Ritenbaugh, J. O. Hill, L. L. Birch, L. D. Frank, K. Glanz, D. A. Himmelgreen, M. Mudd, B. M. Popkin, K. A. Rickard, S. St Jeor, and N. P. Hays. 2001. Environmental and societal factors affect food choice and physical activity: Rationale, influences, and leverage points. *Nutrition Reviews* 59(3, Supplement 2):S21-S39.

Bowen, S., and A. B. Zwi. 2005. Pathways to "evidence-informed" policy and practice: A framework for action. *PLoS Medicine* 2(7):600-605.

Campbell, D. T., and D. A. Kenny. 1999. *A primer on regression artifacts*. New York: Guilford.

Campbell, D. T., and J. C. Stanley. 1966. *Experimental and quasi-experimental designs for research*. Chicago, IL: Rand McNally.

Cargo, M., and S. L. Mercer. 2008. The value and challenges of participatory research: Strengthening its practice. *Annual Review of Public Health* 29:325-350.

CDC (Centers for Disease Control and Prevention). 2009a. *Nutrition resources for health professionals*. http://www.cdc.gov/nutrition/professionals/researchtopractice/index.html (accessed January 4, 2010).

CDC. 2009b. *Racial and ethnic approaches to community health*. http://www.cdc.gov/reach/index.htm (accessed January 4, 2010).

Chatfield, C. 2004. *The analysis of time series: An introduction*. 6th ed. Boca Raton, FL: Chapman & Hall/CRC Press.

Cook, T. D. 2008. "Waiting for life to arrive": A history of the regression-discontinuity design in psychology, statistics and economics. *Journal of Econometrics* 142(2):636-654.

Dobbins, M., S. Jack, H. Thomas, and A. Kothari. 2007. Public health decision-makers' informational needs and preferences for receiving research evidence. *Worldviews on Evidence-Based Nursing* 4(3):156-163.

Dobrow, M. J., V. Goel, and R. E. G. Upshur. 2004. Evidence-based health policy: Context and utilisation. *Social Science and Medicine* 58(1):207-217.

Dobrow, M. J., V. Goel, L. Lemieux-Charles, and N. A. Black. 2006. The impact of context on evidence utilization: A framework for expert groups developing health policy recommendations. *Social Science and Medicine* 63(7):1811-1824.

D'Onofrio, C. N. 2001. Pooling information about prior interventions: A new program planning tool. In *Handbook of program development for health behavior research and practice*, edited by S. Sussman. Thousand Oaks, CA: Sage Publications. Pp. 158-203.

Economos, C. D., R. R. Hyatt, J. P. Goldberg, A. Must, E. N. Naumova, J. J. Collins, and M. E. Nelson. 2007. A community intervention reduces BMI z-score in children: Shape Up Somerville first year results. *Obesity* 15(5):1325-1336.

French, S. A., M. Story, and R. W. Jeffery. 2001. Environmental influences on eating and physical activity. *Annual Review of Public Health* 22:309-335.

Goldberg, J. P., J. J. Collins, S. C. Folta, M. J. McLarney, C. Kozower, J. Kuder, V. Clark, and C. D. Economos. 2009. Retooling food service for early elementary school students in Somerville, Massachusetts: The Shape Up Somerville experience. *Preventing Chronic Disease* 6(3).

Green, L. W. 2006. Public health asks of systems science: To advance our evidence-based practice, can you help us get more practice-based evidence? *American Journal of Public Health* 96(3):406-409.

Green, L. W. 2008. Making research relevant: If it is an evidence-based practice, where's the practice-based evidence? *Family Practice* 25 (Supplement 1):i20-i24.

Green, L. W., and R. E. Glasgow. 2006. Evaluating the relevance, generalization, and applicability of research: Issues in external validation and translation methodology. *Evaluation & the Health Professions* 29(1):126-153.

Green, L. W., and M. W. Kreuter. 2005. *Health program planning: An educational and ecological approach*. 4th ed. New York: McGraw-Hill Higher Education.

Greenhalgh, T., E. Kristjansson, and V. Robinson. 2007. Realist review to understand the efficacy of school feeding programmes. *British Medical Journal* 335(7626):858-861.

Hahn, J., P. Todd, and W. Van Der Klaauw. 2001. Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica* 69(1):201-209.

Hennigan, K. M., M. L. Del Rosario, L. Heath, T. D. Cook, J. D. Wharton, and B. J. Calder. 1982. Impact of the introduction of television on crime in the United States: Empirical findings and theoretical implications. *Journal of Personality and Social Psychology* 42(3):461-477.

Horowitz, C. R., M. Robinson, and S. Seifer. 2009. Community-based participatory research from the margin to the mainstream: Are researchers prepared? *Circulation* 119(19):2633-2642.

Imbens, G. W., and T. Lemieux. 2008. Regression discontinuity designs: A guide to practice. *Journal of Econometrics* 142(2):615-635.

IOM (Institute of Medicine). 2007. *Progress in preventing childhood obesity: How do we measure up?* Edited by J. Koplan, C. T. Liverman, V. I. Kraak, and S. L. Wisham. Washington, DC: The National Academies Press.

Khuder, S. A., S. Milz, T. Jordan, J. Price, K. Silvestri, and P. Butler. 2007. The impact of a smoking ban on hospital admissions for coronary heart disease. *Preventive Medicine* 45(1):3-8.

Kreuter, M. W. 1992. PATCH: Its origin, basic concepts, and links to contemporary public health policy. *Journal of Health Education* 23(3):135-139.

Lightwood, J. D., D. Collins, H. Lapsley, and T. E. Novotny. 2000. Estimating costs of tobacco use. In *Tobacco control in developing countries*, edited by P. Jha and F. J. Chaloupka. Oxford: Oxford University Press. Pp. 63-103.

Ludwig, J., and D. L. Miller. 2007. Does Head Start improve children's life chances? Evidence from a regression discontinuity design. *Quarterly Journal of Economics* 122(1):159-208.

Mays, N., C. Pope, and J. Popay. 2005. Systematically reviewing qualitative and quantitative evidence to inform management and policy-making in the health field. *Journal of Health Services Research and Policy* 10(Supplement 1):6-20.

McLeroy, K. R., D. Bibeau, A. Steckler, and K. Glanz. 1988. An ecological perspective on health promotion programs. *Health Education Quarterly* 15(4):351-377.

Minkler, M., and N. Wallerstein, eds. 2008. *Community-based participatory research for health: From process to outcomes*. 2nd ed. San Francisco: Jossey-Bass.

National Committee on Vital and Health Statistics. 2002. *Shaping a vision of health statistics for the 21st century.* Hyattsville, MD: National Center for Health Statistics.

NCI and SAMHSA (National Cancer Institute and the Substance Abuse and Mental Health Services Administration). 2009. *Research-tested intervention programs*. http://rtips.cancer.gov/rtips/index.do (accessed January 4, 2010).

Ogilvie, D., M. Egan, V. Hamilton, and M. Petticrew. 2005. Systematic reviews of health effects of social interventions: 2. Best available evidence: How low should you go? *Journal of Epidemiology and Community Health* 59(10):886-892.

Ottoson, J. M., and L. W. Green. 1987. Reconciling concept and context: Theory of implementation. *Advances in health education and promotion* 2:353-382.

Ottoson, J. M., L. W. Green, W. L. Beery, S. K. Senter, C. L. Cahill, D. C. Pearson, H. P. Greenwald, R. Hamre, and L. Leviton. 2009. Policy-contribution assessment and field-building analysis of the Robert Wood Johnson Foundation's Active Living Research Program. *American Journal of Preventive Medicine* 36(2, Supplement):S34-S43.

Parcel, G. S. 1987. *Smoking control among women: A CDC community intervention handbook*. Atlanta, GA: Centers for Disease Control, Center for Health Promotion and Education, Division of Health Education.

Reichardt, C. S. 2006. The principle of parallelism in the design of studies to estimate treatment effects. *Psychological Methods* 11(1):1-18.

Reynolds, K. D., and S. G. West. 1987. A multiplist strategy for strengthening nonequivalent control group designs. *Evaluation Review* 11(6):691-714.

Richard, L., L. Potvin, N. Kishchuk, H. Prlic, and L. W. Green. 1996. Assessment of the integration of the ecological approach in health promotion programs. *American Journal of Health Promotion* 10(4):318-328.

Roos, Jr., L. L., N. P. Roos, and P. D. Henteleff. 1978. Assessing the impact of tonsillectomies. *Medical Care* 16(6):502-518.

Rosenbaum, P. R. 2002. *Observational studies*. 2nd ed. New York: Springer.

Rosenbaum, P. R., and D. B. Rubin. 1983. The central role of the propensity score in observational studies for causal effects. *Biometrika* 70(1):41-55.

Rubin, D. B. 1977. Assignment to treatment group on the basis of a covariate. *Journal of Educational Statistics* 2(1):1-26.

Rubin, D. B. 2006. *Matched sampling for causal effects*. New York: Cambridge.

Sallis, J. F., N. Owen, and E. B. Fisher. 2008. Ecological models of health behavior. In *Health behavior and health education: theory, research, and practice*. 4th ed, edited by K. Glanz, B. K. Rimer and K. Viswanath. San Francisco, CA: Jossey-Bass.

Schaalma, H., and G. Kok. 2006. A school HIV-prevention program in the Netherlands. In *Planning health promotion programs: An intervention mapping approach*, edited by L. K. Bartholomew, G. S. Parcel, G. Kok, and N. H. Gottleib. San Francisco, CA: Jossey-Bass. Pp. 511-544.

Schafer, J. L., and J. Kang. 2008. Average causal effects from nonrandomized studies: A practical guide and simulated example. *Psychological Methods* 13(4):279-313.

Shadish, W. R., and T. D. Cook. 1999. Design rules: More steps toward a complete theory of quasi-experimentation. *Statistical Science* 14:294-300.

Shadish, W. R., T. D. Cook, and D. T. Campbell. 2002. *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin.

Simons-Morton, B. G., G. S. Parcel, N. M. O'Hara, S. N. Blair, and R. R. Pate. 1988a. Health-related physical fitness in childhood: Status and recommendations. *Annual Review of Public Health* 9:403-425.

Simons-Morton, B. G., G. S. Parcel, and N. M. O'Hara. 1988b. Implementing organizational changes to promote healthful diet and physical activity at school. *Health Education Quarterly* 15(1):115-130.

Simons-Morton, B. G., W. H. Green, and N. H. Gottleib. 1995. *Introduction to health education and health promotion*. 2nd ed. Prospect Heights, IL: Waveland Press.

Simons-Morton, D. G., B. G. Simons-Morton, G. S. Parcel, and J. F. Bunker. 1988a. Influencing personal and environmental conditions for community health: A multilevel intervention model. *Family and Community Health* 11(2):25-35.

Simons-Morton, D. G., S. G. Brink, G. S. Parcel, K. Tiernan, C. M. Harvey, and J. M. Longoria. 1988b. *Promoting physical actvity among adults: A CDC community intervention handbook*. Atlanta, GA: Centers for Disease Control.

Trochim, W. 1984. *Research design for program evaluation: The regression-discontinuity approach*. Newbury Park, CA: Sage.

West, S. G., and F. Thoemmes. 2010. Campbell's and Rubin's perspectives on causal inference. *Psychological Methods* 15(1):18-37.

West, S., J. Biesanz, and S. Pitts. 2000. Causal inference and generalization in field settings: Experimental and quasi-experimental designs. In *Handbook of research methods in social and personality psychology*, edited by H. Reis and C. Judd. Cambridge: University Press. Pp. 40-84.

West, S., J. Duan, W. Pequegnat, P. Gaist, D. Des Jarlais, D. Holtgrave, J. Szapocznik, M. Fishbein, B. Rapkin, M. Clatts, and P. Mullen. 2008. Alternatives to the randomized controlled trial. *American Journal of Public Health* 98(8):1359-1366.

Wu, W., S. G. West, and J. N. Hughes. 2008a. Effect of retention in first grade on children's achievement trajectories over 4 years: A piecewise growth analysis using propensity score matching. *Journal of Educational Psychology* 100(4):727-740.

Wu, W., S. G. West, and J. N. Hughes. 2008b. Short-term effects of grade retention on the growth rate of Woodcock-Johnson III broad math and reading scores. *Journal of School Psychology* 46(1):85-105.

F

Agendas from Two Workshops

AGENDA

Workshop on Meeting the Challenges of Generating Useful Evidence and Using It Effectively in Obesity Prevention Decision Making

January 8, 2009
The Keck Center
Washington, DC

Goal: Gain insights from experts on how an evidence framework could be helpful in generating and using evidence for decision making about obesity prevention and other multifaceted areas of health promotion and disease prevention.

PROGRAM

Welcoming Remarks
Moderator: *Shiriki Kumanyika, PhD, MPH,* Committee Chair

Panel I: Generating and Evaluating Evidence: Alternatives and Trade-offs

Moderator: *Larry Green, DrPH,* Committee Member

Key Questions:

- Do null studies or insufficient evidence mean that prevention doesn't work?
- What determines whether a prevention study "fails" or "succeeds"? What can we learn from evidence generation and evaluation in other disciplines?
- How can systems thinking inform evidence generation and evaluation?

June Stevens, PhD, RD, University of North Carolina at Chapel Hill
J. Michael Oakes, PhD, University of Minnesota

Madhabi Chatterji, PhD, Columbia University
David Abrams, PhD, American Legacy Foundation

Panel II: Current Approaches to Using Evidence in Obesity Prevention and Related Areas of Health Promotion and Disease Prevention

Moderator: *Ross Brownson, PhD,* Committee Member

Key Question: What are the current challenges and limitations in the application of evidence to the development of guidelines and practice standards in public health efforts?

Joseph Hagan, Jr., MD, University of Vermont
David L. Katz, MD, MPH, Yale University
Stephen Daniels, MD, PhD, University of Colorado Denver School of Medicine, The Children's Hospital

Panel III: Interface of Evidence and Public Policy

Moderator: *Adolfo Valadez, MD, MPH,* Committee Member

Key questions:

- Where does evidence fit into public policy decision making?
- How do policy makers gather, use, and evaluate evidence in decision making?
- What challenges do policy makers face in finding and using evidence?
- What types of evidence are most useful to policy makers?

Joshua Sharfstein, MD, Baltimore City Health Department
Stephanie Coursey Bailey, MD, MS, Office of Public Health Practice, Centers for Disease Control and Prevention
Lynn Silver, MD, MPH, New York City Department of Health and Mental Hygiene

CLOSING REMARKS

AGENDA

Workshop on the Application of Systems Thinking to the Development and Use of Evidence in Obesity Prevention Decision Making

March 16, 2009
Beckman Center Auditorium
Irvine, California

Goal: Gain insights from experts on how a systems science framework and modeling could be helpful in generating and using evidence for decision making about obesity prevention and other multifaceted areas of health promotion and disease prevention.

PROGRAM

Welcoming Remarks
Shiriki Kumanyika, PhD, MPH, Committee Chair

UK Foresight Programme's Tackling Obesities: Future Choices Project
Susan Jebb, PhD (via videoconference), Medical Research Council Collaborative Center for Human Nutrition Research, Cambridge, UK

Knowledge Synthesis and Integration: Changing Models, Changing Practices
Allan Best, PhD, Vancouver Coastal Health Research Institute

System Dynamics Simulation in Support of Obesity Prevention Decision Making
Bobby Milstein, PhD, MPH, Centers for Disease Control and Prevention
Jack Homer, PhD, Homer Consulting

Multilevel Approaches to Understanding and Preventing Obesity: Analytical Challenges and New Directions
Ana Diez Roux, MD, PhD, MPH, University of Michigan School of Public Health

CLOSING REMARKS

G

Committee Member Biographical Sketches

Shiriki K. Kumanyika, Ph.D., M.S.W., M.P.H., R.D. (*Chair*), is Professor of Epidemiology in the Department of Biostatistics and Epidemiology and Pediatrics (Gastroenterology, Nutrition Section) and Associate Dean for Health Promotion and Disease Prevention at the University of Pennsylvania School of Medicine. Dr. Kumanyika has a unique interdisciplinary background that integrates epidemiology, nutrition, prevention, minority health, and women's health issues across the life course. She received a B.A. from Syracuse University, an M.S.W. from Columbia University, a Ph.D. in human nutrition from Cornell University, and an M.P.H. from The Johns Hopkins University. The main themes in Dr. Kumanyika's research concern the role of nutritional factors in the primary and secondary prevention of chronic diseases, with a particular focus on obesity, sodium reduction, and related health problems such as hypertension and diabetes. She has a particular interest in the epidemiology and management of obesity among African Americans. Dr. Kumanyika has served on numerous national and international advisory committees and expert panels related to nutrition and obesity, and currently serves as Vice Chair of the Department of Health and Human Services' (HHS') Secretary's Advisory Committee for Healthy People 2020. She is co-chair of the International Obesity Task Force and serves as a consultant to the World Health Organization's Department of Nutrition for Health and Development. She has served on the Food and Nutrition Board of the Institute of Medicine (IOM) and a number of IOM committees, including the Committee on Prevention of Obesity in Children and Youth and the Committee on Progress in Preventing Childhood Obesity (as a volunteer consultant). Dr. Kumanyika is a member of the IOM.

David B. Abrams, Ph.D., is Executive Director of the Steven A. Schroeder Institute for Tobacco Research and Policy Studies at the American Legacy Foundation and is Professor of Health, Behavior, and Society at The Johns Hopkins University Bloomberg School of Public Health. He earned a Ph.D. from Rutgers University in clinical psychology. Dr. Abrams is a licensed clinical psychologist, specializing in health psychology, addictions, and tobacco use behavior. His primary interest is in fostering the systems integration of biomedical, sociobehavioral, and ecologic-public health models to understand and eliminate tobacco use behavior at the population level. Dr. Abrams was Director of the Office of Behavioral and Social Sciences Research (OBSSR) at the National Institutes of Health (NIH), from 2005 to 2008. Prior to that, he was Professor of Psychiatry and Human Behavior and Professor of Community Health at Brown University Medical School and founding Director of Brown's Transdisciplinary Centers for Behavioral and Preventive Medicine for more than 16 years. He has published more than 250 scholarly articles and has been a principal or co-investigator on 65 grant awards, including a National Cancer Institute Program Project award for Transdisciplinary Tobacco Use Research Centers of Excellence. Dr. Abrams is lead author of *The Tobacco Dependence Treatment Handbook: A Guide to Best Practices* and recipient of a book of the year award from the *American Journal of Nursing*. He is a Fellow and past President of the Society of Behavioral Medicine and a recipient of the society's Distinguished Scientist Award, Distinguished Service Award, and Distinguished Research Mentor Award. He served on the IOM Committee on Reducing Tobacco Use: Strategies, Barriers, and Consequences.

Ross C. Brownson, Ph.D., is Professor of Epidemiology at Washington University in St. Louis. He is involved in numerous community-level studies designed to reduce modifiable risk factors such as physical inactivity, unhealthful eating, and tobacco use. In particular, he is interested in the impacts of environmental and policy interventions on health behaviors. Dr. Brownson also conducts research on dissemination of evidence-based interventions in public health settings. His research is supported by NIH, the Centers for Disease Control and Prevention (CDC), the Missouri Foundation for Health, and the Robert Wood Johnson Foundation. He is Co-Director of the CDC-funded Prevention Research Center (PRC)—a 5-year, multimillion dollar project aimed at developing innovative approaches to chronic disease prevention. The PRC in St. Louis is a joint research endeavor between the Saint Louis University School of Public Health and Washington University. Dr. Brownson is the author of six books and more than 250 peer-reviewed articles. His books include *Chronic Disease Epidemiology and Control, Applied Epidemiology, Handbook of Obesity Prevention, Communicating Public Health Information Effectively: A Guide for Practitioners*, and *Evidence-Based Public Health*. He is Associate Editor of the *Annual Review of Public Health* and on the editorial board of three other journals. He is a former and founding member of the 15-person CDC task force responsible for developing the *Guide to Community Preventive Services*.

Frank Chaloupka, Ph.D., is Distinguished Professor in the Department of Economics at the University of Illinois at Chicago (UIC); he also holds an appointment in the UIC School of Public Health's Division of Health Policy and Administration and is director of the UIC Health Policy Center. He received his Ph.D. from the City University of New York Graduate School and University Center, specializing in health economics. Dr. Chaloupka's research interests include the effects of prices, policies, and other environmental factors on health behavior, including cigarette smoking and other tobacco use, alcohol use and abuse, illicit drug use, diet, and physical activity, as well as outcomes related to these behaviors. Much of his research focuses on youth and young adults. He received a University Scholar Award from the University of Illinois for his research on the economic analysis of substance use and abuse and was an expert reviewer on the IOM's National Cancer Policy Board in 1997.

Madhabi Chatterji, Ph.D., is Associate Professor of Measurement and Evaluation and Director of the Assessment and Evaluation Research Initiative at Teachers College, Columbia University. Dr. Chatterji received her Ph.D. in measurement, evaluation, and research methodology from the University of South Florida, and has taught methodology courses at the graduate and advanced graduate levels for more than 15 years. Her research interests are in assessment and evaluation methodology, evidence standards and new evidence synthesis approaches, educational equity, and standards-based educational reforms in the United States. Her research has received national and regional recognition through awards from the American Educational Research Association in 2004 (AERA-Division H, 2004—Advances in Research Methodology) and the Florida Educational Research Association (Distinguished Paper Award, 1993). As a recent Fulbright Research Scholar (2007-2008), Dr. Chatterji is examining issues related to universal education and gender equality at the primary level in two South Asian nations, India and Bangladesh. She has published widely on both measurement and evaluation topics, with several papers weighing the utility of systemic, mixed-method designs against randomized controlled trials for studying complex education and public health interventions. Dr. Chatterji is a frequent consultant in her areas of specialty and serves on several national and international advisory panels in education, public health, and medicine. She took part in the planning meeting for this committee's work.

Barbara A. Dennison, M.D., is Director of Policy and Research Translation in the Division of Chronic Disease and Injury Prevention at the New York State Department of Health, where she has overall responsibility for identifying and working with programs to ensure the implementation of evidence-based, best-practice, and promising approaches to preventing and reducing risk factors and conditions related to chronic diseases and injury. Dr. Dennison serves as the point of contact for the Division of Chronic Disease and Injury Prevention in developing policy and legislative initiatives to improve public health, and as the division's policy liaison with other Department of

Health units, state agencies, and external organizations in the development of chronic disease prevention policies. From 2003 to 2009, she served as Director, Bureau of Health Risk Reduction, where she provided leadership for the state's community-based prevention efforts to reduce obesity, heart disease, stroke, diabetes, and other chronic diseases and to improve guideline-concordant care of chronic diseases.

Christina Economos, Ph.D., is Associate Professor of Nutrition and New Balance Chair in Childhood Nutrition at the Friedman School of Nutrition Science and Policy at Tufts University. Dr. Economos earned her M.S. at Columbia University and her Ph.D. at the Friedman School. Her research efforts focus on the interactions among exercise, diet, and body composition. Her translational research includes theory-based obesity prevention interventions with ethnically and socioeconomically diverse children, adolescents, and their families in urban and rural communities across the United States. She was principal investigator for the Shape Up Somerville (SUS) project, funded primarily by CDC, and currently leads several large obesity prevention intervention trials. The SUS project targeted behavior change in children through community-based, environmental change in a low-income, racially/ethnically diverse population. Results published in *Obesity* in 2007 demonstrate that this multipronged intervention reduced overweight in the treatment children compared with children in two similar comparison communities. Dr. Economos has held positions in public health nutrition, including a tenure at the Massachusetts Department of Public Health. She serves on numerous state and national advisory boards. She was a consultant on the Youth Subcommittee for the 2008 Physical Activity Guidelines and is a member of the Public Information Council of the American Society for Nutrition.

Steven Gortmaker, Ph.D., is Professor of the Practice of Health Sociology in the Department of Society, Human Development, and Health and Director of the Health Prevention Research Center at the Harvard School of Public Health. He received a Ph.D. from the University of Wisconsin, Madison. Dr. Gortmaker's research is focused on the health of children and adolescents, particularly households living in poverty and minority populations. The major goal of this research has been to identify modifiable risks for morbidity and mortality in the young, and to both initiate and evaluate interventions to improve these outcomes. Dr. Gortmaker's work has focused on a broad variety of risks, ranging from sociological concepts such as poverty, social stress, and social networks to behaviors such as smoking, inactivity (exemplified by television viewing), and diet, and on interventions including policy at the national and state levels and programs at the regional, county, school, hospital, clinic, and individual levels. His research encompasses collaborative work with research groups at Harvard, in the Boston area, nationally, and internationally. Dr. Gortmaker also was involved in two school-based interventions: the Eat Well and Keep Moving Project and Planet Health. Both programs have been shown to reduce obesity risk behaviors in experimental trials and are in wide use across the United States.

Lawrence W. Green, Dr.P.H., is Co-Director of the Society, Diversity, and Disparities Program for the Cancer Center and Professor in the Department of Epidemiology and Biostatistics at the University of California, San Francisco (UCSF). Prior to his current position, Dr. Green was Director of the Office of Science and Extramural Research, Public Health Practice Office, at CDC. He received a B.S., M.P.H., and Dr.P.H., each in the area of public health or public health education, from the University of California, Berkeley. The majority of his work focuses on developing and strengthening practice-based evidence as a complement to evidence-based practices derived from systematic reviews of highly controlled trials. In the Carter Administration, he served as the first Director of the Office of Health Information, Health Promotion, Physical Fitness and Sports Medicine. He has served on the public health faculties at Berkeley, Johns Hopkins, Harvard, the University of Texas, and the University of British Columbia and as Vice President and Director of the Kaiser Family Foundation's national Health Promotion Program. Dr. Green is a member of the IOM and was a member ex officio of the IOM's Clinical Research Roundtable from 2001 to 2005.

Robert A. Hiatt, M.D., Ph.D., is Professor and Co-Chair of the Department of Epidemiology and Biostatistics at UCSF and Director of Population Science and Deputy Director of the UCSF Helen Diller Family Comprehensive Cancer Center. He is also an Adjunct Professor, Division of Epidemiology, at the University of California, Berkeley. From 1998 to early 2003, he was the first Deputy Director of the Division of Cancer Control and Population Sciences at the National Cancer Institute, where he oversaw cancer research in epidemiology and genetics, surveillance, and health services. His research interests have included the epidemiology of cancer, especially breast cancer; cancer prevention and screening; health services and outcomes research; and the social determinants of cancer. He is a recipient of NIH Merit Awards for the Quality of Cancer Care Committee and the Cancer Progress Report, and also received an NIH Director's Award for the latter. Dr. Hiatt has served on two IOM committees in the past, one on improving cancer care in Georgia and the other on NIH research strategies to eliminate health disparities.

William Purcell, III, J.D., former Mayor of Nashville, Tennessee, is now Director of the Institute of Politics at the Kennedy School of Government at Harvard University. Mr. Purcell graduated from Hamilton College and Vanderbilt University School of Law. In 1986 he was elected to the Tennessee House of Representatives, where he served for five terms, assuming the positions of Majority Leader and Chair of the Select Committee on Children and Youth. After retiring from the General Assembly, he founded and became Director of the Child and Family Policy Center at the Vanderbilt Institute of Public Policy Studies, which focuses on collaboration among academic research, politics, and best practices. During his tenure as Mayor of Nashville, the city saw a more than 50 percent increase in school funding and the creation of more than 20,000 affordable housing units. In 2006, *Governing Magazine* named him Public Official of the Year.

Robert Sege, M.D., Ph.D., is Chief of Ambulatory Pediatrics at the Boston Medical Center and Professor of Pediatrics at Boston University School of Medicine. Dr. Sege received his medical degree from Harvard Medical School and his Ph.D. in biology from the Massachusetts Institute of Technology. He is a primary care pediatrician whose research involves the development of a health care response to violence involving children and adolescents, ranging from child abuse to youth violence. For the past 2 years, Dr. Sege has led an effort to develop a new framework for making evidence-based policy recommendations for child health promotion activities in the clinical setting. He is a member of the American Academy of Pediatrics Evidence Working Group. Previously, Dr. Sege directed the federally funded development of the American Academy of Pediatrics' violence prevention program, Connected Kids: Safe, Strong, Secure, and was co-investigator in a large prospective national study of how primary care providers identify and manage child physical abuse. He is principal investigator of Project Dulce, a child maltreatment research and demonstration project.

Harold Sox, M.D., M.A.C.P., is Professor of Medicine and of The Dartmouth Institute (emeritus) at Dartmouth Medical School and immediate past Editor of the *Annals of Internal Medicine*. Dr. Sox received his medical degree from Harvard Medical School. He began his studies in medical decision making when he was Chief Medical Resident at Dartmouth Medical School, where he later became department Chair. In his lifetime of research and writing, he has explored such issues as technology assessment, medical decision making, disease prevention and health promotion, cost-effectiveness analysis, physicians' and patients' risk preferences, and medical education. He has served on the editorial boards of other medical publications, such as *Scientific American Medicine* and *The New England Journal of Medicine*, and is the author of *Medical Decision-Making* and other books. Dr. Sox has led national committees that have shaped clinical, educational, and public policy in the United States. He has served as Chair of the U.S. Preventive Services Task Force, the IOM Committee to Study HIV Transmission through Blood Products, and the IOM Committee on Health Effects of Exposures in the Persian Gulf War, and was founding Chair of the Medicare Coverage Advisory Committee. Dr. Sox has served as President of the American College of Physicians and is a member of the IOM.

Adolfo M. Valadez, M.D., M.P.H., is Assistant Commissioner for Prevention and Preparedness Services at the Texas Department of Health Services. Dr. Valadez received his medical degree from the University of Texas Medical Branch at Galveston and is board-certified in internal medicine. Later, he completed the Commonwealth Fund/Harvard University Fellowship in Minority Health Policy and received an M.P.H. degree from the Harvard School of Public Health. Dr. Valdez oversees approximately 1,000 employees and a budget of more than $450 million. His three primary responsibilities are overseeing disease control and prevention, disaster preparedness and response, and the agency's laboratory services. Previously, Dr. Valadez served as Medical Director and Health Authority for the Austin/Travis County Health and Human Services Department. In the past, he also served as Medical Director of the Martha Eliot Health Center in Jamaica Plain, Massachusetts, and as a primary care provider. Dr. Valadez has served on peer review committees for the Health Resources and Services Administration's Centers of Excellence grants, as a member of a CDC Expert Panel on Physical Activity and Nutrition and a Special Emphasis Panel on community-based participatory research methodologies, and on numerous professional societies' advisory committees.

Leticia Van de Putte, R.Ph., is a Senator in the Texas State Senate, currently in her fourth term, after having served as a State Representative for five terms. Senator Van de Putte holds a bachelor's degree in pharmacy from the University of Texas at Austin. As Senator, she has focused on issues regarding children, veterans, improved access to health care, quality education, and economic development. She has consistently authored and sponsored bills to assist families in securing opportunities, and has also been a key author of legislation reforming the welfare system, revamping the Juvenile Code, and insuring children. In 1993, she was a Kellogg Fellow at the Kennedy School of Government at Harvard University. Senator Van de Putte has received awards from numerous civic organizations and community groups. She has been actively involved in the National Conference of State Legislatures, serving as President from 2006 to 2007. In 2002, she became Chair of the Senate Democratic Caucus. In 2008, she served as Co-Chair of the Democratic National Convention.

Stephen G. West, Ph.D., is Professor in the Department of Psychology at Arizona State University and Visiting Professor, Free University of Berlin, Germany (summer semesters). He received his Ph.D. from the University of Texas at Austin. He is past Editor of *Psychological Methods* and current Associate Editor of *Multivariate Behavioral Research*. His quantitative research interests include experimental, quasi-experimental, and longitudinal designs for field research, structural equations modeling, multiple regression analysis, mediational analysis, multilevel analysis, longitudinal data analysis, and graphics and exploratory data analysis. His substantive research interests focus on personality research and prevention-related issues in health, mental health, and substance abuse. He recently completed work on grant-funded projects involving a randomized trial of an intervention for couples at high risk for HIV/AIDS and a longitudinal study of the well-being of HIV/AIDS patients. He is currently working on grant-funded projects that examine short- and long-term effects of grade retention on achievement and on personality, cognitive, and social development in children; the development and validation of English and Spanish language asthma screening questionnaires for inner city communities; and statistical mediation models for prevention research.

Index

D

E

F

G

H

I

N

O

P

Q

R

S

T

U

V

W